

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:29 : Search time 18 Seconds

(without alignments)
644.035 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRSLSIFGATGSGESTFPL.....MDHLARRAEERAGLRQOKR 394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	741.5	37.4	399	4	US-09-491-362-7
2	741.5	37.4	399	4	US-09-874-562-7
3	741.5	37.4	477	4	US-09-449-335-6
4	737.5	37.2	477	4	US-09-449-335-6
5	737.5	37.2	477	4	US-09-480-921B-29
6	736	37.1	477	4	US-09-480-921B-10
7	654.5	33.0	475	4	US-09-491-362-2
8	654.5	33.0	475	4	US-09-874-562-2
9	138.5	7.0	2293	4	US-09-368-590-2
10	115	5.8	5215	4	US-09-105-537-2
11	113.5	5.7	7257	3	US-09-335-409-5
12	113.5	5.7	7257	4	US-09-568-102-5
13	113.5	5.7	7257	4	US-09-567-968-5
14	113.5	5.7	7257	4	US-09-568-480-5
15	113.5	5.7	7257	4	US-09-568-486-5
16	113.5	5.7	7257	4	US-09-567-899-5
17	113.5	5.7	7257	4	US-09-567-899-5
18	111	5.6	5087	4	US-09-144-085-1
19	110.5	5.6	3729	2	US-08-804-227C-4
20	109.5	5.5	4472	2	US-08-804-227C-2
21	108.5	5.5	3724	2	US-08-804-227C-10
22	108.5	5.5	3724	2	US-08-804-198-4
23	107.5	5.4	6095	4	US-09-144-085-2
24	106.5	5.4	609	4	US-08-969-683A-67
25	103.5	5.2	4551	3	US-09-320-878-1
26	103.5	5.2	4613	3	US-09-105-537-31
27	103.5	5.2	11877	4	US-09-105-537-6

28	102	5.1	1346	3	US-09-320-878-4	Sequence 4, Appl1
29	102	5.1	1346	4	US-09-105-537-37	Sequence 37, Appl1
30	102	5.1	1580	2	US-08-804-227C-11	Sequence 11, Appl1
31	102	5.1	1580	2	US-08-804-198-5	Sequence 5, Appl1
32	99.5	5.0	3170	4	US-09-036-987A-4	Sequence 4, Appl1
33	99.5	5.0	3170	4	US-09-370-700-4	Sequence 4, Appl1
34	98.5	5.0	419	1	US-08-305-172B-4	Sequence 4, Appl1
35	98	4.9	1382	3	US-09-057-570-4	Sequence 4, Appl1
36	97.5	4.9	3739	3	US-09-320-878-2	Sequence 2, Appl1
37	97.5	4.9	3739	4	US-09-105-537-33	Sequence 33, Appl1
38	97	4.9	3170	2	US-07-642-734C-5	Sequence 5, Appl1
39	97	4.9	3170	3	US-08-923-009A-5	Sequence 5, Appl1
40	96.5	4.9	523	2	US-08-923-287-2	Sequence 2, Appl1
41	96.5	4.9	523	4	US-09-385-287-2	Sequence 2, Appl1
42	96.5	4.9	2152	4	US-09-036-987A-3	Sequence 3, Appl1
43	96.5	4.9	2152	4	US-09-370-700-3	Sequence 3, Appl1
44	95.5	4.8	1864	2	US-08-804-227C-3	Sequence 3, Appl1
45	95.5	4.8	3519	4	US-09-428-517-4	Sequence 4, Appl1

ALIGNMENTS

```
RESULT 1
US-09-491-362-7
Sequence 7, Application US/09491362
Patent No. 6281017
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Lange, Bernd M
TITLE OF INVENTION: 1-DEOXY-D-XYULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
FILE REFERENCE: MSURIA977
CURRENT APPLICATION NUMBER: US/09/491,362
CURRENT FILING DATE: 2000-01-26
EARLIER APPLICATION NUMBER: 60/118,349
EARLIER FILING DATE: 1999-02-03
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 399
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-491-362-7

Query Match      37.4% Score 741.5 DB 4: Length 399:
Best Local Similarity 41.3% Pred. No. 1.5e-67:
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6:

QY 2 RSTISFGATGSGESTFPLVMKRGPEAFRTAALTGNIRRLAEKARLAEVTAHE 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 3 KRISIVGSGISGTQGLDVAE--NDKRFVVALAGSVITLLADQVRFKALAVRRE 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 DCLPALREALAGTG--TEVAGGAQAIATAEADRP-ADWTMSAIVGAAGLVGRALKHGRT 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 SLINKEKELADLDYKLEITIEGQVIEVARHDEAVTVTGIVCGALPTVAALIEAGD 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LALANKESTVITAGOLMRTAENGATILPVDSEHSVFOALGEDTACGERVITRASGP 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 IALANKESTVITAGOLMRTAENGATILPVDSEHSVFOALGEDTACGERVITRASGP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 FRDWSLERTACTVAEAQAHPNMWQORISISASMFNALELIETREFGEPRIEAV 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 FRDWPPEKLEKVAADALHPNMWQORISISASMFNALELIETREFGEPRIEAV 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 VHPDSIVHAAGVCGDGLMAHLGPADMRHAIGPALNMPGR--GEVPAVIDLAOTASIT 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 IHPDSIVHMTODSSVLAOLGMPDMRPIIYTMSMPDRVPCSEVTWPRDLCKLSIT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 FKKPDERPALRLARDVAAAGLSAANFAKETAIDHFTIGRIGFLDMAVVEETL-- 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 FKKPDERPALRLARDVAAAGLSAANFAKETAIDHFTIGRIGFLDMAVVEETL-- 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

QY 354 --AGVSTDPLEFGKVPDALEEVYLAADHILARRAA 383
Db 361 HRNELVTSP-----SLEEIVHYDLMAAREYA 385

RESULT 2

US-09-874-562-7
; Sequence 7, Application US/09874562
; Patent No. 6420159
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: MSUR17549
; CURRENT APPLICATION NUMBER: US/09/874,562
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/491,362
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/118,349
; PRIOR FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-874-562-7

Query Match 37.4%; Score 741.5; DB 4; Length 399;
Best Local Similarity 41.3%; Pred. No. 1.5e-67;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGTSTPDLVNRKGGPEAFRTVALTGCNIRRLAEMARALKAEIAYTAHE 61
Db 3 KRISIVGTSIGTQTDIVAE--NPDKFRVYALAGSNVTLADQVRRFPALVAVRNE 60
QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP--ADWTMSAIVGAGLVPGMRALKHGR 118
Db 61 SLINELKEALADLDYKLEIIPGEGVIEVARHPRPAVTVGTCAGIKPTVAIAEAGKD 120
QY 119 LALANKESLYTAGOLIMRTAENGATILPVDESHSAVQALAGEDTACVEYVITASGCP 178
Db 121 IALANKETLLAGPVPVPLARKHNKXILIPADSEHSALFOCIQIPEGALRKIIITASGCA 180
QY 179 FRDMSLERIRACTVAEAOAHNPMWNGORISIDSASMEKALELJETREFGFEEDRIEAY 238
Db 181 FRDMPVERKLVKAYKADALKHNNMNGKKITVDSATLFRKGLVEIAHLFCAEYDDIEIV 240
QY 239 VHPOSIYHAMVGFCDGGLMAHLGPAIDNRHAIQFALNMPGR--GEVYARIDLAQIASLT 295
Db 241 IHPOSIHSMIETODSSVLAQLGWPMDMRLPILYTMSPDRVPCSEVYWPRLDCKLSLT 300
QY 296 FOKPDEERFRLRLAROVMAARGISGAFAFNKAEIADHFIAGRTIGFLDMAVVEEYL-- 353
Db 301 FKKPDNKKYPSMDLAIVAAGRGAGTGTGVLASANKAKAEMFIDEKISTYIDIFKVEYELTCDK 360
QY 354 --AGVSTDPLEFGKVPDALEEVYLAADHILARRAA 383
Db 361 HRNELVTSP-----SLEEIVHYDLMAAREYA 385

RESULT 3

US-09-449-335-6
; Sequence 6, Application US/09449335
; Patent No. 6303365
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Heatt, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Kloti, Andreas
; TITLE OF INVENTION: Method of determining the activity of

; TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase

; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/449,335
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: DE 199 35 967.9
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-449-335-6

Query Match 37.4%; Score 741.5; DB 4; Length 477;
Best Local Similarity 41.3%; Pred. No. 2e-67;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGTSTPDLVNRKGGPEAFRTVALTGCNIRRLAEMARALKAEIAYTAHE 61
Db 81 KRISIVGTSIGTQTDIVAE--NPDKFRVYALAGSNVTLADQVRRFPALVAVRNE 138
QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP--ADWTMSAIVGAGLVPGMRALKHGR 118
Db 139 SLINELKEALADLDYKLEIIPGEGVIEVARHPRPAVTVGTCAGIKPTVAIAEAGKD 198
QY 119 LALANKESLYTAGOLIMRTAENGATILPVDESHSAVQALAGEDTACVEYVITASGCP 178
Db 199 IALANKETLLAGPVPVPLARKHNKXILIPADSEHSALFOCIQIPEGALRKIIITASGCA 258
QY 179 FRDMSLERIRACTVAEAOAHNPMWNGORISIDSASMEKALELJETREFGFEEDRIEAY 238
Db 259 FRDMPVERKLVKAYKADALKHNNMNGKKITVDSATLFRKGLVEIAHLFCAEYDDIEIV 318
QY 239 VHPOSIYHAMVGFCDGGLMAHLGPAIDNRHAIQFALNMPGR--GEVYARIDLAQIASLT 295
Db 319 IHPOSIHSMIETODSSVLAQLGWPMDMRLPILYTMSPDRVPCSEVYWPRLDCKLSLT 378
QY 296 FOKPDEERFRLRLAROVMAARGISGAFAFNKAEIADHFIAGRTIGFLDMAVVEEYL-- 353
Db 379 FKKPDNKKYPSMDLAIVAAGRGAGTGTGVLASANKAKAEMFIDEKISTYIDIFKVEYELTCDK 438
QY 354 --AGVSTDPLEFGKVPDALEEVYLAADHILARRAA 383
Db 439 HRNELVTSP-----SLEEIVHYDLMAAREYA 463

RESULT 4

US-09-449-335-2
; Sequence 2, Application US/09449335
; Patent No. 6303365
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Kloti, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/449,335
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: DE 199 35 967.9
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-449-335-2

[illegible]

```
Db 319 IHPOSIHSMETQDSSVLAQLGMPDMRLPLLYTMSWPDRCSEVTPRDLCKLSLT 378
Oy 296 FOKDEERFPALRLARVMAARGLSGAAPNAKKEIALDHFIFAGSLGFLDMAVAVEETL-- 353
Db 379 FKADNWKYSMDLAYAAGRAAGTGTGVLNANKEAVEMFIDEXISYIDLIFKVELTCDK 438
Oy 354 --AGVSTDPLEFGKVPDALEEVLAHDHLARRAAE---AAGLR 390
Db 439 HRNELVTS-----SLEELVHYDLMAREYADVOJSSGAR 473

RESULT 7
US-09-491-362-2
; Sequence 2, Application US/09491362
; Patent No. 6281017
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLOLOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; FILE REFERENCE: USUR14977
; CURRENT FILING DATE: 2000-01-26
; EARLIER APPLICATION NUMBER: 60/118,349
; EARLIER FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mentha piperita
US-09-491-362-2

Query Match
Best Local Similarity 33.0%; Score 654.5; DB 4; Length 475;
Matches 156; Conservative 73; Mismatches 152; Indels 17; Gaps 9;

Oy 2 RSLIFGATGSGISTDLYMRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHE 61
Db 80 KPIVISIGTSGIGTQTLDIYAE--NPKFRIVALAAGSNVTLADQ-KAFKPKLVSVKDE 136
Oy 62 DCLPALREALAG--TGTEVAGCAQAIAEADRP-ADWTMSAIVGAGLVPGMRALKHGR 118
Db 137 SLISELEALAGFEDMPDRIIPGEGMIEVARHPDAVTVTGIVCAGLKPVAALIEAGKD 196
Oy 119 LALANKESLYTAGOLKMTAEOENGATILPVDESHAVFOALAGEDTACVEVITTAAGP 178
Db 197 IALANKETLLAGGFVPLAKKHNVKILLPADSEHSAIFOCIGLPEGLRRIITLASGA 256
Oy 179 FRDMSLERIRACTVAEAOAHNMGMGORISIDSASM-FNKALELITREFGFEPPDRIEA 237
Db 257 FRDLPVEKLKVKYADALKHSNMGMKNTVRLQLFENKLEVIKAYHLEGAEDDIEI 316
Oy 238 VVHPOSTIVHNAVGFCDGLMAHLGPAIDMRHAIQFALNMPGR---GEVVARIDLAQIASL 294
Db 317 VHSPTSIHSMVETQDSSVLAQLGMPDMRLPLLYTMSPEVYVSEITWPRDLCKV-DL 375
Oy 295 TFOKDERFPALRLARVMAARGLSGAAPNAKKEIALDHFIFAGSLGFLDMAVAVEETIA 354
Db 376 PFKPKDREIRPAMDLAYAAMKSRSTMGTGVLNANKEAVEMFIDEXISYIDLIFKVELTCD 435
Oy 355 GVSTDPLEFGKVPDALEEVLAHDHLARRAAE---BAAGL 389
Db 436 KHRSEM---AVSPSLEELVHYDQWARYDAATVLKSAGL 470

RESULT 8
US-09-674-562-2
; Sequence 2, Application US/09874562
; Patent No. 6420159
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
```

```
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLOLOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; FILE REFERENCE: USUR17549
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/491,362
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/118,349
; PRIOR FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mentha piperita
US-09-674-562-2

Query Match
Best Local Similarity 33.0%; Score 654.5; DB 4; Length 475;
Matches 156; Conservative 73; Mismatches 152; Indels 17; Gaps 9;

Oy 2 RSLIFGATGSGISTDLYMRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHE 61
Db 80 KPIVISIGTSGIGTQTLDIYAE--NPKFRIVALAAGSNVTLADQ-KAFKPKLVSVKDE 136
Oy 62 DCLPALREALAG--TGTEVAGCAQAIAEADRP-ADWTMSAIVGAGLVPGMRALKHGR 118
Db 137 SLISELEALAGFEDMPDRIIPGEGMIEVARHPDAVTVTGIVCAGLKPVAALIEAGKD 196
Oy 119 LALANKESLYTAGOLKMTAEOENGATILPVDESHAVFOALAGEDTACVEVITTAAGP 178
Db 197 IALANKETLLAGGFVPLAKKHNVKILLPADSEHSAIFOCIGLPEGLRRIITLASGA 256
Oy 179 FRDMSLERIRACTVAEAOAHNMGMGORISIDSASM-FNKALELITREFGFEPPDRIEA 237
Db 257 FRDLPVEKLKVKYADALKHSNMGMKNTVRLQLFENKLEVIKAYHLEGAEDDIEI 316
Oy 238 VVHPOSTIVHNAVGFCDGLMAHLGPAIDMRHAIQFALNMPGR---GEVVARIDLAQIASL 294
Db 317 VHSPTSIHSMVETQDSSVLAQLGMPDMRLPLLYTMSPEVYVSEITWPRDLCKV-DL 375
Oy 295 TFOKDERFPALRLARVMAARGLSGAAPNAKKEIALDHFIFAGSLGFLDMAVAVEETIA 354
Db 376 PFKPKDREIRPAMDLAYAAMKSRSTMGTGVLNANKEAVEMFIDEXISYIDLIFKVELTCD 435
Oy 355 GVSTDPLEFGKVPDALEEVLAHDHLARRAAE---BAAGL 389
Db 436 KHRSEM---AVSPSLEELVHYDQWARYDAATVLKSAGL 470

RESULT 9
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2

Query Match
7.0%; Score 138.5; DB 4; Length 2293;
```


	Query Match	5.8%;	Score 115;	DB 4;	Length 5215;	
	Best Local Similarity	20.2%;	Pred. No. 0.11;			
	Matches 108;	Conservative	52;	Mismatches 162;	Indels 212;	Gaps 21.
QY	28 EAFRTVLTGGRNIRRLAEMARALKAELA-----VTANEDCPALREALAGTGEVAG	80				
	: : : :					
Db	1509 QAVRLFTEFLDPRDVTVADLGRSLAAFTAEFKKALTATRELLAGDALP-GRGQATG	1567				
QY	81 -----GAQAIA-----EAARDPMWTMSATV	101				
	:					
Db	1568 LVTPGERAPARGTAFLTGOGAKORVAMGBELRAHNFVAALDTYYUAAIDRHILDRPLREIV	1627				
QY	102 GAA-----GLVP-----GMRLKH-GRTLLANKE	125				
Db	1628 AAGEEIDLTAAYTORALFAFEVALFRLLHNGLGVLPDLITGSVSGEILAAAHAHVLSLDNA	1687				
QY	126 SLVYTAGQLMRTAOENGATILRYDSEHSNVFALAGEDTACERYIITASGGRF-----	179				
	: : : :					
Db	1668 RLVTATRGRLIMOSARKGG-MIAVQAGEEVEVESLGYE---GRAVAANAHPNPTAVVVG	1742				
	: : : :					

[illegible]

Db 5778 DRLSV-----AVNSPRSTVLAGE-PAALAEVLAI 5806

RESULT 12

US-09-568-102-5

Sequence 5, Application US/09568102

Patent No. 6346404

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyf, Devon

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/568,102

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-568-102-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

QY 4 LSIFGATGSGESTFDLVRKRGPEAFVTVALTGGRNIRRLAEMARALKAEIA-----VT 58

Db 5510 VSSFGVSG-----TNAHVLEAPAAAFAPAAARSALFVLSAKSAALDAQARLSAHV 5565

QY 59 AHEDC-LPALREALAGCTEV-----AGGAQAIAEADRPADMTWSAIVGAGVPGM 110

Db 5566 AHPGLGDLAFSLATRSPTVRLAVALTSREALSALD-----TAAQQAAPPA 5615

QY 111 RALKHGRFLANKESLYTAGQLMRTAENGATILPDSHSAVFOALAGED----- 163

Db 5616 AARGHAST-GSAPKVVFVFPQ-----GSGWLMGOKLSEEVFPRDL 5658

QY 164 TACVRYVITTAGSGPRFDMSLERIRACTVAEAQHPNMSGQRISIDSASFNKALELIE 223

Db 5659 SAC-DRAIQAEAG-----WSL-----LAELADETTSQGRIDVQPALEAEVALSA 5705

QY 224 TREFFGEPRRIEAVHPQSIVHAMVGCDCGLMAHLGPRADMRHAIGFALMPGRGEVPV 283

Db 5706 LMRSMGVEPPDAV-----VGHSM-----GEVAA 5727

QY 284 ARIDLAQIASLTFQKPEERFPALRLARDVMAARGLSGAAPNAKETAIDHFIAGRIGFL 343

Db 5728 AHV--AGALSL-----EDAVAI-ICRSLRLRISGCGMAVVELSLAEAEALLGYE 5777

QY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVLA 375

Db 5778 DRLSV-----AVNSPRSTVLAGE-PAALAEVLAI 5806

RESULT 13

US-09-567-969-5

Sequence 5, Application US/09567969

Patent No. 6355457

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyf, Devon

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/567,969

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-567-969-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

QY 4 LSIFGATGSGESTFDLVRKRGPEAFVTVALTGGRNIRRLAEMARALKAEIA-----VT 58

Db 5510 VSSFGVSG-----TNAHVLEAPAAAFAPAAARSALFVLSAKSAALDAQARLSAHV 5565

QY 59 AHEDC-LPALREALAGCTEV-----AGGAQAIAEADRPADMTWSAIVGAGVPGM 110

Db 5566 AHPGLGDLAFSLATRSPTVRLAVALTSREALSALD-----TAAQQAAPPA 5615

QY 111 RALKHGRFLANKESLYTAGQLMRTAENGATILPDSHSAVFOALAGED----- 163

Db 5616 AARGHAST-GSAPKVVFVFPQ-----GSGWLMGOKLSEEVFPRDL 5658

QY 164 TACVRYVITTAGSGPRFDMSLERIRACTVAEAQHPNMSGQRISIDSASFNKALELIE 223

Db 5659 SAC-DRAIQAEAG-----WSL-----LAELADETTSQGRIDVQPALEAEVALSA 5705

QY 224 TREFFGEPRRIEAVHPQSIVHAMVGCDCGLMAHLGPRADMRHAIGFALMPGRGEVPV 283

Db 5706 LMRSMGVEPPDAV-----VGHSM-----GEVAA 5727

QY 284 ARIDLAQIASLTFQKPEERFPALRLARDVMAARGLSGAAPNAKETAIDHFIAGRIGFL 343

Db 5728 AHV--AGALSL-----EDAVAI-ICRSLRLRISGCGMAVVELSLAEAEALLGYE 5777

QY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVLA 375

Db 5778 DRLSV-----AVNSPRSTVLAGE-PAALAEVLAI 5806

RESULT 14

US-09-568-480-5

Sequence 5, Application US/09568480

Patent No. 6355458

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyf, Devon

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/568,480

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-568-480-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Best Local Similarity 24.0%; Pred. No. 0.26;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

```
OY 4 LSIFGATSGISGTFDLMRKGPPEAFRTVALTGRNIRRLAEMARLAKAELA-----VT 58
Db 5510 VSSFGVSG-----TNAHVLEEPAPAAAFAPAAARSAEFLVLSAKSAALDAQAARLSAHV 5565
OY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADWTMSAIVGACGLVPGM 110
Db 5566 AHPELGIDDLAFSLATTSPTMYRLAANAATSRALSAALD-----TAAGQAPPA 5615
OY 111 RALKHGRLLALANKESLVTAGQLMRTAENGATLLPVDSHSAVFOALAGED----- 163
Db 5616 AARGHAST-GSAPKVVFFEPGQ-----GSQWLGNGOKLLSEEPFRDAL 5658
OY 164 TACVERVITITASGGPFRRDMSLERIACVTAEOAHPPNMSGORISIDSASFENKALELIE 223
Db 5659 SAC-DRAIQAEAG-----WST-----LAELADETTSQGRIDVVPALFAIEVALSA 5705
OY 224 TREFFGEPPDRIEAVVHPQSIYHAWGFCGGLMAHLGPADMRHAIGFALMMPGRGEVPV 283
Db 5706 LMRSGVPEPDV-----VGHSM-----GEVAA 5727
OY 284 ARIDLAQIASLTFOKPDERFPALRLADVMAARGLSGAFAFNAKEIALDHFIAGRIGFL 343
Db 5728 AHV--AGALSL-----EDAVAI-ICRRSLILRRISGOGEMAVVELSLAEAEALLGYE 5777
OY 344 DMAAVEETLAGVSTD-----LFGKVPDALEEVILAM 375
Db 5778 DRLSY-----AVSNSPRSTVLAGE-PALAEVLAI 5806
```

RESULT 15

```
US-09-568-486-5
; Sequence 5, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligou, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zitzke, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-5
```

Query Match 5.7%; Score 113.5; DB 4; Length 7257;
Best Local Similarity 24.0%; Pred. No. 0.26;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

```
OY 4 LSIFGATSGISGTFDLMRKGPPEAFRTVALTGRNIRRLAEMARLAKAELA-----VT 58
Db 5510 VSSFGVSG-----TNAHVLEEPAPAAAFAPAAARSAEFLVLSAKSAALDAQAARLSAHV 5565
OY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADWTMSAIVGACGLVPGM 110
Db 5566 AHPELGIDDLAFSLATTSPTMYRLAANAATSRALSAALD-----TAAGQAPPA 5615
OY 111 RALKHGRLLALANKESLVTAGQLMRTAENGATLLPVDSHSAVFOALAGED----- 163
Db 5616 AARGHAST-GSAPKVVFFEPGQ-----GSQWLGNGOKLLSEEPFRDAL 5658
```

```
OY 164 TACVERVITITASGGPFRRDMSLERIACVTAEOAHPPNMSGORISIDSASFENKALELIE 223
Db 5659 SAC-DRAIQAEAG-----WST-----LAELADETTSQGRIDVVPALFAIEVALSA 5705
OY 224 TREFFGEPPDRIEAVVHPQSIYHAWGFCGGLMAHLGPADMRHAIGFALMMPGRGEVPV 283
Db 5706 LMRSGVPEPDV-----VGHSM-----GEVAA 5727
OY 284 ARIDLAQIASLTFOKPDERFPALRLADVMAARGLSGAFAFNAKEIALDHFIAGRIGFL 343
Db 5728 AHV--AGALSL-----EDAVAI-ICRRSLILRRISGOGEMAVVELSLAEAEALLGYE 5777
OY 344 DMAAVEETLAGVSTD-----LFGKVPDALEEVILAM 375
Db 5778 DRLSY-----AVSNSPRSTVLAGE-PALAEVLAI 5806
```

Search completed: April 15, 2003, 14:00:29
Job time : 37 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:25 ; Search time 37 Seconds
(without alignments)
1418.939 Million cell updates/sec

```

Title: US-09-673-198-30
Perfect score: 1984
Sequence: 1 MRSLSIFGATGSGESTFDL.....MDHLARRAEAGLRQQR 394

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470.

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1984	100.0	394	20	AAV52839	Rhodobacter sphaer
2	918.5	46.3	386	23	AAO21875	Isoprenoid related
3	876	44.2	388	23	AAO21879	Isoprenoid related
4	876	44.2	388	23	AAO21883	Isoprenoid related
5	778	39.2	394	23	ABG61582	High growth methan
6	778	39.2	394	23	AAE22302	Methylomonas 16a
7	778	39.2	394	23	AAU80326	Methylomonas 16a
8	775.5	39.1	396	23	AAO21879	Isoprenoid related
9	756.5	38.1	473	21	AAAB01369	1-deoxy-D-xylinose
10	746	37.6	472	21	AAAB01367	1-deoxy-D-xylinose

11	743.5	37.5	394	23	AAO21886	Isoprenoid related
12	742	37.4	473	21	AAO21832	1-deoxy-D-xylulose
13	741.5	37.4	399	21	AAO97157	A. thaliana parlia
14	741.5	37.4	406	21	AAO21810	A. thaliana DXPri
15	741.5	37.4	406	21	AAO21370	1-deoxy-D-xylulose
16	741.5	37.4	420	21	AAO21413	A. thaliana DXPri
17	741.5	37.4	476	21	AAO30294	Arbidopsin D-1-de
18	741.5	37.4	477	22	AAO21966	1-deoxy-D-xylulose
19	741.5	37.4	477	23	AAO21922	Isoprenoid related
20	741.5	37.4	477	23	AAO39322	Herbicidally activ
21	737.5	37.2	477	22	AAO31965	1-deoxy-D-xylulose
22	736	37.1	477	21	AAO95658	Arbidopsin herbic
23	732	36.9	398	23	AAO21878	Isoprenoid related
24	732	36.9	398	23	AAO21885	Isoprenoid related
25	725.5	36.6	475	21	AAO21388	1-deoxy-D-xylulose
26	720.5	36.3	477	21	AAO95659	Arbidopsin herbic
27	717	36.1	397	23	AAO21877	Isoprenoid related
28	717	36.1	397	23	AAO21887	Isoprenoid related
29	717	36.1	397	23	AAO90946	H. influenzae DXR
30	716	36.1	398	20	AAO52838	Escherichia coli p
31	716	36.1	400	23	AAO48689	Isoprenoid related
32	702.5	35.4	380	23	AAO48625	Listeria monocytog
33	695.5	35.1	394	23	AAO21881	Isoprenoid related
34	695.5	35.1	394	23	AAO21888	Isoprenoid related
35	689.5	34.8	499	21	AAO21363	1-deoxy-D-xylulose
36	685	34.5	402	23	AAO21880	Isoprenoid related
37	685	34.5	402	23	AAO21890	Isoprenoid related
38	670	33.8	388	23	AAO21876	Isoprenoid related
39	670	33.8	388	23	AAO21892	Isoprenoid related
40	665.5	33.5	386	23	AAO21891	Isoprenoid related
41	654.5	33.0	475	21	AAO97156	Menha piperita 1-
42	654.5	33.0	475	21	AAO21371	1-deoxy-D-xylulose
43	644	33.5	411	22	AAO56274	Propionibacterium
44	628	31.7	433	23	AAO21894	Isoprenoid related
45	628	31.7	436	22	AAO21207	Mycobacterium tube

ALIGNMENTS

XX	RESULT 1
XX	AAY52839
XX	ID AAY52839 standard; Protein; 394 AA.
XX	
AC	AAY52839;
XX	
DT	26-JAN-2000 (first entry)
DE	Rhodobacter sphaeroides protein sequence SEQ ID NO:30.
XX	
KW	Isoprenoid; microorganism; detection; antibacterial; herbicide;
KW	heart disease; osteoporosis; haemostasis; cancer; immunopotentialion;
KW	health food; antifouling coating; larnesyl pyrrolinic acid; pyruvic acid;
KW	1-deoxy-D-xylulose-5-phosphate; glyceraldehyde-3-phosphate;
XX	2-C-methyl-D-erythritol-4-phosphate.
XX	
OS	Rhodobacter sphaeroides.
XX	
PN	W09953071-A1.
PD	
PD	21-OCT-1999.
XX	
PE	14-APR-1999; 99WO-JP01987.
XX	
PR	14-APR-1998; 98JP-0103101.
PR	05-AUG-1998; 98JP-0221910.
PR	15-FEB-1999; 99JP-0035739.
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Miyake K, Hashimoto S, Motoyama H, Ozaki A, Seto H, Kuzuyama T;
PI	Takahashi S;
XX	

DR WP1: 1999-620434/53.
DR N-PSDB: AAZ33164.
PT Preparation of recombinant isoprenoid compounds useful for treatment of
PT heart diseases, osteoporosis and hemostasis, preventing cancer and
PT immunopotentialiation
PS Claim 6; Page 130-132; 145pp; Japanese.
XX
XX The present invention describes the preparation of an isoprenoid
XX compound comprising using at least 1 DNA e.g. encoding proteins which
XX elevate the efficiency of the synthesis or DNA encoding a farnesyl
XX pyrolytic producing enzyme. The method of preparation of an isoprenoid
XX compound comprises using at least 1 DNA, a vector, cloned cells, their
XX derived recombinant DNAs or transformed products in a culture system and
XX extracting the isoprenoid accumulated in the medium. The DNA encodes at
XX least 1 of the following: (a) a compound for activating or catalyzing
XX the production of 1-deoxy-D-xylose-5-phosphate from pyruvic acid and
XX glyceraldehyde-3-phosphate; (b) an enzyme producing farnesyl pyrolytic
XX acid; (c) a protein which elevates the efficiency of synthesis of
XX isoprenoid compounds and comprises a 3 or 4 amino acid sequence
XX optionally with 1 or more of the amino acids being deleted or
XX substituted or an additional amino acid being inserted; (d) a protein
XX which activates or catalyzes the production of 2-C-methyl-D-erythritol-
XX 4-phosphate from 1-deoxy-D-xylose-5-phosphate; or (e) a protein which
XX activates a target compound or reaction and is a string end or hybrid of
XX the DNA encoded in (a)-(e). Isoprenoid compounds are useful in drugs
XX (e.g. for the treatment of heart diseases, osteoporosis and hemostasis,
XX for preventing cancer and as immunopotentialiators), health foods and
XX antifouling coatings. The isoprenoids also inhibit enzymatic reactions
XX on the non-mevalonate pathway and can be used as antibacterials and
XX herbicides. The present sequence is used in the exemplification of
XX the present invention.
SQ Sequence 394 AA;
Query Match 100.0%; Score 1984; DB 20; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.8e-193;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSLSIFGATGSGESTFDLMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
Db 1 MRSLSIFGATGSGESTFDLMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
QY 61 EDCPLPALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
Db 61 EDCPLPALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
QY 121 LANESLVTAGQLMRTAENGATILPVDSEHSVFOALAGEEDTACVERVITITASGGPFR 180
Db 121 LANESLVTAGQLMRTAENGATILPVDSEHSVFOALAGEEDTACVERVITITASGGPFR 180
QY 181 DWSLERIRACTVAEAQAHPNMSKORISIDSASFNKALELIEFREFFGFPDRIEAVVH 240
Db 181 DWSLERIRACTVAEAQAHPNMSKORISIDSASFNKALELIEFREFFGFPDRIEAVVH 240
QY 241 POSTIVHAMVGFCDGGLMAHLGIPADMRHAIGFALMWPGRGEVVARIDLAQIASLTFOKPD 300
Db 241 POSTIVHAMVGFCDGGLMAHLGIPADMRHAIGFALMWPGRGEVVARIDLAQIASLTFOKPD 300
QY 301 EERPPALRLADVAARGLSGAANNAKEIALDHFIAGRIGFLDMAAVEETLAGVSTDP 360
Db 301 EERPPALRLADVAARGLSGAANNAKEIALDHFIAGRIGFLDMAAVEETLAGVSTDP 360
QY 361 LFGKVPDALEEVLAHDLARRAAEEAAGLRQOKR 394
Db 361 LFGKVPDALEEVLAHDLARRAAEEAAGLRQOKR 394
RESULT 2
AAO21875
ID AAO21875 standard; Protein: 386 AA.
XX

AC AAO21875;
XX
XX 13-SEP-2002 (first entry)
DT
XX
DE Isoprenoid related protein sequence SEQ ID No 97.
XX
XX Isoprenoid; CoQ(10); 1-deoxyxylylucose-5-phosphate synthase; DXS; DDS;
XX decaprenyl diphosphate synthase.
XX Spilngomonas treperei.
XX WO200226933-A2.
XX 04-APR-2002.
XX 28-SEP-2001; 2001WO-US30328.
XX 29-SEP-2000; 2000US-236580P.
XX (CRGI) CARGILL INC.
XX Gokarn R, Jessen H, Zidwick MJ;
XX WP1: 2002-416480/44.
XX Substantially pure polypeptides having e.g.,
XX 1-deoxyxylylucose-5-phosphate synthase activity, useful for the
XX production of isoprenoids, especially CoQ(10)
XX
XX Disclosure; Fig 30; 246pp; English.
XX
XX The invention relates to methods and materials for the production of
XX isoprenoids. More particularly the invention provides isolated nucleic
XX acids, substantially pure polypeptides, host cells, and methods for
XX producing various isoprenoid compounds. The polypeptides are useful for
XX the production of isoprenoids, especially CoQ(10). Expressing the pure
XX polypeptides, which has 1-deoxyxylylucose-5-phosphate synthase (DXS)
XX activity or decaprenyl diphosphate synthase (DDS) activity, is useful for
XX increasing production of CoQ(10) in a cell having endogenous DDS
XX activity. This sequence represents a protein relating to the isoprenoid
XX production of the invention.
SQ Sequence 386 AA;
Query Match 46.3%; Score 918.5; DB 23; Length 386;
Best Local Similarity 50.4%; Pred. No. 8e-85;
Matches 194; Conservative 62; Mismatches 122; Indels 7; Gaps 3;
QY 1 MRSLSIFGATGSGESTFDLMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
Db 2 VKRVTVLGATGSGVSTLTLIER--NPHAFEVALTANCDDVEKLAAAIIRTRACAVVAD 59
QY 61 EDCPLPALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
Db 61 EDCPLPALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
QY 121 LANESLVTAGQLMRTAENGATILPVDSEHSVFOALAGEEDTACVERVITITASGGPFR 180
Db 121 LANESLVTAGQLMRTAENGATILPVDSEHSVFOALAGEEDTACVERVITITASGGPFR 180
QY 181 DWSLERIRACTVAEAQAHPNMSKORISIDSASFNKALELIEFREFFGFPDRIEAVVH 240
Db 181 DWSLERIRACTVAEAQAHPNMSKORISIDSASFNKALELIEFREFFGFPDRIEAVVH 240
QY 241 POSTIVHAMVGFCDGGLMAHLGIPADMRHAIGFALMWPGRGEVVARIDLAQIASLTFOKPD 300
Db 241 POSTIVHAMVGFCDGGLMAHLGIPADMRHAIGFALMWPGRGEVVARIDLAQIASLTFOKPD 300
QY 301 EERPPALRLADVAARGLSGAANNAKEIALDHFIAGRIGFLDMAAVEETLAGVSTDP 360
Db 301 EERPPALRLADVAARGLSGAANNAKEIALDHFIAGRIGFLDMAAVEETLAGVSTDP 360
QY 361 LFGKVPDALEEVLAHDLARRAAEEAAGLRQOKR 394
Db 361 LFGKVPDALEEVLAHDLARRAAEEAAGLRQOKR 394

Db 358 ---AABETLDAVLDAEALYAAE 379

RESULT 3

ID AAO21879 standard; Protein; 388 AA.

AC AAO21879;

DT 13-SEP-2002 (first entry)

DE Isoprenoid related protein sequence SEQ ID No 101.

XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;
KW decaprenyl diphosphate synthase.

OS Zymonas mobilis.

PN WO200226933-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30328.

PR 29-SEP-2000; 2000US-236580P.

PA (CRGI) CARGILL INC.

PI Gokarn R, Jessen H, Zidwick MJ;

DR WPI; 2002-416480/44.

XX Substantially pure polypeptides having e.g.,
PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
PI production of isoprenoids, especially CoQ(10)

XX Disclosure; Fig 27; 246pp; English.

XX The invention relates to methods and materials for the production of
CC isoprenoids. More particularly the invention provides isolated nucleic
CC acids, substantially pure polypeptides, host cells, and methods for
CC producing various isoprenoid compounds. The polypeptides are useful for
CC the production of isoprenoids, especially CoQ(10). Expressing the pure
CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)
CC activity or decaprenyl diphosphate synthase (DDS) activity, is useful for
CC increasing production of CoQ(10) in a cell having endogenous DDS
CC activity. This sequence represents a protein relating to the isoprenoid
CC production of the invention.

XX Sequence 388 AA;

Query Match 44.2%; Score 876; DB 23; Length 388;

Best Local Similarity 46.9%; Pred. No. 1.7e-80; Indels 10; Gaps 3;

Matches 183; Conservative 67; Mismatches 130;

2 RLSITFGATSGESTFEDLVMRKSGEAPRTVALTGRNIRRLAEMARKALAYTAHE 61

5 RVTYVAGATSGISHSLDIER--NLDROYALITANRWKDLADAKRTNKRVIADP 62

62 DCLPALREALAGTGVAGGAQAIAEADRPADMTSAIVGAAGLVGMRALKHGRTLAL 121

63 SLYNDLKEALAGSSVEAAGADALVEAAMGADMTMAIIGCAGLKATLAIIRKGTVAL 122

122 ANKESLVTAAGQLMRTAENGATILFVDESHSAVFQALAGEDTACVERVIITASGGPFD 181

123 ANKESLVSAAGGLMIDAVREHGTLLPVDSEHNATFOCPHHNDYVRRIITASGGPFD 182

182 WSLERIRACTVAEAQAHPNWSMGORISIDSMFNKALELIETREFGEPDRIEAVVHP 241

183 TSLAEATVATYPERAVOHPNWSMGAKISIDSATMKNKGLEIEVYHLPFPLEKFELTVHP 242

242 QSLVHAMVGCDCGGLMAHLGAPADMRAHIGFALNWPGRGEVPAVIDLAQIASLTFOKPDE 301

Db 243 QSYIHSWVEYLDGSIQAQSGPDMRTPIGHTLAMPFRMETPASEIDFTLRQMDPEADPY 302

302 EREPALRLARDYMAANGLSGAFAFNAAKETALDHFHFGIFLDMAAVVEETLAGVSTDPL 361

303 EREPALTLAMESIKSGAPAVMANAEIYAVAFDCKIGFLDIKIVERTL-----DHY 357

362 FGKVPDALEEVLAAMDHLARRAEAGLRQ 391

358 TPATPSSLEDVFAIDNEAR---IQAAALME 384

RESULT 4

ID AAO21883 standard; Protein; 388 AA.

AC AAO21883;

DT 13-SEP-2002 (first entry)

DE Isoprenoid related protein sequence SEQ ID No 116.

XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;
KW decaprenyl diphosphate synthase.

OS Zymonas mobilis.

PN WO200226933-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30328.

PR 29-SEP-2000; 2000US-236580P.

PA (CRGI) CARGILL INC.

PI Gokarn R, Jessen H, Zidwick MJ;

DR WPI; 2002-416480/44.

XX Substantially pure polypeptides having e.g.,
PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
PI production of isoprenoids, especially CoQ(10)

XX Disclosure; Fig 32; 246pp; English.

XX The invention relates to methods and materials for the production of
CC isoprenoids. More particularly the invention provides isolated nucleic
CC acids, substantially pure polypeptides, host cells, and methods for
CC producing various isoprenoid compounds. The polypeptides are useful for
CC the production of isoprenoids, especially CoQ(10). Expressing the pure
CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)
CC activity or decaprenyl diphosphate synthase (DDS) activity, is useful for
CC increasing production of CoQ(10) in a cell having endogenous DDS
CC activity. This sequence represents a protein relating to the isoprenoid
CC production of the invention.

XX Sequence 388 AA;

Query Match 44.2%; Score 876; DB 23; Length 388;

Best Local Similarity 46.9%; Pred. No. 1.7e-80; Indels 10; Gaps 3;

Matches 183; Conservative 67; Mismatches 130;

2 RLSITFGATSGESTFEDLVMRKSGEAPRTVALTGRNIRRLAEMARKALAYTAHE 61

5 RVTYVAGATSGISHSLDIER--NLDROYALITANRWKDLADAKRTNKRVIADP 62

62 DCLPALREALAGTGVAGGAQAIAEADRPADMTSAIVGAAGLVGMRALKHGRTLAL 121

63 SLYNDLKEALAGSSVEAAGADALVEAAMGADMTMAIIGCAGLKATLAIIRKGTVAL 122

122 ANKESLVTAAGQLMRTAENGATILFVDESHSAVFQALAGEDTACVERVIITASGGPFD 181

```

D- 123 ANKESLVSAGLMDAVAREHGTLLPVDSEHNAIFOCPPHNRDVRRIITTAGSGPRT 182
QY 182 WSLERIRACTYAEAOAHNMGMGORISIDSASFNKALELETFEEFEEDREAVVHP 241
D- 183 TSLAEKATVTPERAVOHNMGMGAKISIDSATMANKLELEVAHLEFPILEKEEILVHP 242
QY 242 QSIYHAWGFCDDGLMAHLGPADNRHAIIGFALNMPGRGEVVARIDLAQIASLTFORPDE 301
D- 243 QSVIHSWVEYLDGSLAQIGSPDMRTPIGHTLAMPKREMTAEESIDFTKLROMFEADY 302
QY 302 ERFPALRLARVMAARGISGAFAFNAKEIALDHFIAGRIGLDMAAVVEETLACVSDPL 361
D- 303 ERFPALRLAMESIKSGGARPAVMNNAEMIAVAFLDKKIGLIDIAKIVETL-----DHX 357
QY 362 FGKVPDALEEVLDHMLARRAEAEAGLRQ 391
D- 358 TPATPSSLEDVFAIDNEAR---TQAAALME 384

RESULT 5
ABG61582
ID ABG61582 standard; protein; 394 AA.
XX
AC ABG61582;
XX
DT 27-AUG-2002 (first entry)
XX
DE High growth methanotrophic bacterial strain: polypeptide #32.
XX
KM High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme:
KM methane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;
KM pyrophosphate dependent phosphofructokinase; nitrogen-containing compound;
KM ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
KM methane-containing environment; waste water treatment system; isoprenoid;
KM nitrous oxide; terpene; animal feed; carotenoid; exopolysaccharide.
XX
OS Methylobionas 16a.
XX
PN WO200220728-A2.
XX
PD 14-MAR-2002.
XX
PE 28-AUG-2001; 2001WO-US26827.
XX
PR 01-SEP-2000; 2000US-229858P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
PI Koffas M, Odom JM, Schenzle A;
XX
DR WPI; 2002-452200/48.
XX
DR N-PSDB; ABK83261.
XX
PT New high growth methanotrophic bacterial strain, useful for producing
PT single cell proteins, grows on a C1 carbon substrate, and comprises a
PT functional gene encoding in Embden-Meyerhof carbon pathway
XX
PS Claim 11; Page 141-142; 157pp; English.
XX
CC The invention relates to a high growth methanotrophic bacterial strain,
CC which grows on a C1 carbon substrate e.g. methane and methanol, and
CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
CC gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16S
CC RNA. The bacterial strain is useful for the production of single cell
CC protein and for the biotransformation of a nitrogen-containing compound,
CC e.g. ammonia, nitrate or nitrogen. It is also useful for the
CC production of a feed product comprising a protein, carbohydrates and a
CC pigment and for reducing oxygen demand, for removing nitrates and
CC nitrates in methane-containing environments such as landfills, waste
CC water treatment systems or anywhere that methane, oxygen and nitrates are
CC present. The bacterial strain of the invention can be used as a
CC denitrifying agent for the conversion of nitrate or nitrite to nitrous

```

```

CC oxide with methane or methanol as a carbon source. It is also used in the
CC production of biomass including proteins, carbohydrates and a wide
CC variety of pigments (particularly for isoprenoid pigments for the
CC purpose of generating animal feeds). In production of terpene and
CC carotenoid compounds, useful as pigments and as monomers in polymeric
CC materials and in production of exopolysaccharides at high levels.
CC Sequences ABG61551-ABG61590 represent high growth methanotrophic
CC bacterial strain proteins of the invention.
XX
SQ Sequence 394 AA;
XX
Query Match 39.28; Score 778; DB 23; Length 394;
Best Local Similarity 41.68; Pred. No. 1.7e-70;
Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;

QY 1 MRSSTIFGATSGISGEFEDLVMRKSGPEAFRTVALTGGRNRLAEAMARLKAELAVTAH 60
D- 1 MKGICITGATSGISGVSTLDVVARH--PDKIOVALTANGNDIALYECGLAHHPYAVVVM 58
QY 61 EDCPLALREALAGT---GTEVAGGAQAIAEAVD--RPADWTSAIVGAAGLPGMRALKHG 116
D- 59 ESKVAEFKQRIASPVADIKVLSGSEALQVATLENDVTMAAIVGAAGLPTLAAGKAG 118
QY 117 RTLALANKESLVTAGQLMRTAQENGATILPVDSHSAVFOAL-----AGEDTACVERV 171
D- 119 KTVLLANKREALYMSGQIFMQAVSDSGAVLLPIDEHNAIFOCMPAGYTPGHTAKQARIL 178
QY 172 ITASGPFPRDMSLERIRACTYAEAOAHNMGMGORISIDSASFNKALELETFEEFGFE 231
D- 179 LTASGPFPRRPIETLSSVTPDOVAHHPKMDGKISVDSATMANKLELETFEACLELME 238
QY 232 PDRIEAVVHPOSIYHAWGFCDDGLMAHLGPADNRHAIIGFALNMPGRGEVVARIDLAQI 291
D- 239 PDQLEAVIHPQSIHSMVDYDGSVLAQMGNDPDMRTPIAHMAMPFERFDSGVAPLDIFEV 298
QY 292 ASLTFQRPDERFPALRLARVMAARGISGAFAFNAKEIALDHFIAGRIGLDMAAVVEE 351
D- 299 GHMDEKRPDLKRFPCRLAIEAISKSGIMPTVLAANEIAEVAFLNEBVEKTDIAVIER 358
QY 352 TLGAVSTDPLEKVPDALEEVLDHMLARRAE 385
D- 359 SMAQFKPD-----AGSLELVLDQDQAREVARD 387

RESULT 6
AAE22302
ID AAE22302 standard; protein; 394 AA.
XX
AC AAE22302;
XX
DT 25-JUL-2002 (first entry)
XX
DE Methylobionas 16a sp. D-1-deoxyxylulose-5-phosphate reductoisomerase.
XX
KM Carotenoid; isopentenyl pyrophosphate; antheraxanthin; astaxanthin; diet;
KM anti-oxidant; steroid; flavour; fragrance; electro-optic application;
KM aquaculture; enzyme; D-1-deoxyxylulose-5-phosphate reductoisomerase; Dxr.
XX
OS Methylobionas 16a sp.
XX
PN WO200218617-A2.
XX
PD 07-MAR-2002.
XX
PE 04-SEP-2001; 2001WO-US27420.
XX
PR 01-SEP-2000; 2000US-229858P.
XX
PR 01-SEP-2000; 2000US-229907P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
PI Brzostowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;
PI Odom JM, Picatagallo SK, Rouviere PE;

```


xy	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																																																							
00	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	10	11	12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	1F	20	21	22	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F	30	31	32	33	34	35	36	37	38	39	3A	3B	3C	3D	3E	3F	40	41	42	43	44	45	46	47	48	49	4A	4B	4C	4D	4E	4F	50	51	52	53	54	55	56	57	58	59	5A	5B	5C	5D	5E	5F	60	61	62	63	64	65	66	67	68	69	6A	6B	6C	6D	6E	6F	70	71	72	73	74	75	76	77	78	79	7A	7B	7C	7D	7E	7F	80	81	82	83	84	85	86	87	88	89	8A	8B	8C	8D	8E	8F	90	91	92	93	94	95	96	97	98	99	00

ID	Sequence	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
359	SAQGFPPD-----AGSELVLQADDAEVAR	387							
DB									
RESULT 8									
ID	AAO21884								
AC	AAO21884 standard; Protein; 396 AA.								
XX									
XX	AAO21884;								
XX									
DT	13-SEP-2002 (first entry)								
XX									
DE	Isoprenoid-related protein sequence SEQ ID No 117.								
XX									
KW	Isoprenoid: CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;								
KM	decaprenyl diphosphate synthase.								
XX									
OS	Pseudomonas aeruginosa.								
XX									
PN	MO200226933-AZ.								
XX									
PD	04-APR-2002.								
XX									
PF	28-SEP-2001; 2001WO-US30328.								
XX									
PR	29-SEP-2000; 2000US-236580P.								
XX									
XX	(CRGI) CARGILL INC.								
PA									
XX	Gokarn R, Jessen H, zidwick MJ;								
PI									
DR	WPI; 2002-416480/44.								
XX									
PT	Substantially pure polypeptides having e.g.,								
PT	1-deoxyxylulose-5-phosphate synthase activity, useful for the								
PT	production of isoprenoids, especially CoQ(10).								
XX									
PS	Disclosure; Fig 32; 246pp; English.								
XX									
CC	The invention relates to methods and materials for the production of								
CC	isoprenoids. More particularly the invention provides isolated nucleic								
CC	acids, substantially pure polypeptides, host cells, and methods for								
CC	producing various isoprenoid compounds. The polypeptides are useful for								
CC	the production of isoprenoids, especially CoQ(10). Expressing the pure								
CC	polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)								
CC	activity or decaprenyl diphosphate synthase (DDS) activity, is useful for								
CC	increasing production of CoQ(10) in a cell having endogenous DDS								
CC	activity. This sequence represents a protein relating to the isoprenoid								
CC	production of the invention.								
XX									
XX									
Sequence	396 AA;								
Query Match	39.1%; Score 775.5; DB 23; Length 396;								
Best Local Similarity	42.8%; Pred. No. 3.2e-70;								
Matches 166; Conservative	73; Mismatches 134; Indels 15; Gaps 5								
4	LSIFGATGSGISTEDPLVMKGGKGEAFRTVALTGGNIRRLAMARALKAELAVTGHEDC 63								
7	ISVLGATGSGISTEDLVORH--PDYREAFALGFSRLAEALCLRHRYAVAVPEQAA 64								
64	IPALREALAGTS--TEVAGAAQAIAEADRP--ADWTMSAIVGAGLVPGMRALKHRTLA 120								
65	ALALGSLAAAGIRIRFVLGELGALCEVAASAPEDVMMAIVGAGLPSTLAANEACKRYL 124</								

```

Db      245 EVVYIHQSVIHSWVDVGDSVAIQGNPMPRTPISYAAMWPRIIDSGVSPLDMFAVRGLD    304
Qy      296 FOKPEDEERPARLRADWMAAGLSGAAPNAKKEIALDLHFIAGRIGFLDMAAVEETLAG    355
        ||::|||::||| |::| : : |::|: : | : : | : : | : : | : : | : : | : : |
Db      305 FORPEORPPCCRLASQAETAESGAPAMLNANVEAAAFLEIRHIFRFSDAVIDEVLNR     364
Qy      356 VSTDPLEFGVPDALAEVLAMOHLEARAA   383
        : : : : : : : : : : | | | :
Db      365 EAVTAV-----ESLDQVLADRRASSVA 387

RESULT 9
AAB01369
ID AAB01369 standard; Protein; 473 AA.
XX
AC AAB01369;
XX
DT 16-NOV-2000 (first entry)
XX
DE 1-deoxy-D-xylulose-5-phosphate reductoisomerase sequence.
XX
KW 1-deoxy-D-xylulose-5-phosphate reductoisomerase; transgenic plant;
KM herbicide; isopentenyl diphosphate; mapping; breeding; phenotype;
KN probe.
XX
OS Triticum aestivum.
XX
PN WO200034448-A1.
PD 15-JUN-2000.
XX
PF 03-DEC-1999; 99WO-US28616.
XX
PR 04-DEC-1998; 98US-0110865.
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Lee J, Tao Y;
XX
DR WPI: 2000-431295/37.
DR N-PSDB: AAA47412.
XX
PT Novel polynucleotide encoding isopentenyl diphosphate biosynthetic
PT enzymes useful for producing transgenic plants with altered isopentenyl
PT diphosphate levels and for selecting polynucleotides affecting
PT expression of the enzyme
XX
PS Claim 2; Page 55-57; 63pp; English.
XX
CC Nucleic acids encoding 1-deoxy-D-xylulose-5-phosphate
CC reductoisomerase are useful for creating transgenic plants with
CC altered levels of plastid isopentenyl diphosphate in cells.
CC The 1-deoxy-D-xylulose 5-phosphate reductoisomerase polypeptides are
CC useful as targets for identifying inhibitors of the enzyme which are
CC useful as herbicides. Nucleic acid fragments of 1-deoxy-D-xylulose-5-
CC phosphate reductoisomerase are useful as probes for genetically and
CC physically mapping genes useful in plant breeding for developing
CC lines with desired phenotypes.
XX
SQ Sequence 473 AA;

Query Match          38.1%; Score 756.5; DB 21; Length 473;
Best Local Similarity 42.3%; Pred.No.3,6e+68;
Matches 166; Conservative 69; Mismatches 138; Indels 19; Gaps 6;

       2 RSLSTFSGATGSIGSETPDLVMRKGPPEARFTVALTGSRNIIRLAEAKALKAELAWTAHE 61
       : ::|||::||| |::| : : |::|: : | : : | : : | : : | : : | : : |
Db      76 KRISIVGSGSIGTOTDIJVAE--NPDKFRVVALLAAGSNVTLLAOVKTKFKKLAVARNR 133

62 DCLPALREALACTG--TEVYGAGAQAIAEAADRP-ADMWTSAIVGAAGLVPGMRALKHGRT 118
       | ::|||::| : : | : : | : : : : ||| ||| : : : : |
Db      134 SLNLNELKEALACGEEMPETIIPEGGVIEVARRHDPADVTVTTGITVGCAGLKPTYAALIEAGND 193

```


CC increasing production of Coq(10) in a cell having endogenous DDS
 CC activity. This sequence represents a protein relating to the isoprenoid
 CC production of the invention.
 XX
 SQ Sequence 394 AA;
 Query Match 37.5%; Score 743.5; DB 23; Length 394;
 Best Local Similarity 43.0%; Pred. No. 5.8e-67;
 Matches 169; Conservative 58; Mismatches 151; Indels 15; Gaps 5;
 YY
 4 LSTIGANGSTGSESTFDLVMRKSGPEAFRTVALTGGRNIRLAEKARALKAEATVAHEDC 63
 D LTLIGSTGSGESTLVDVSRH--PEKFRVFLAGHKQVEKLAACQFFHPYAVVADEH 63
 YY 64 LPALREALA--GTGTEVAGGAQAIAEADR-PADMTSAIVGAAGLVGMRALKHGRTLA 120
 D 64 AARLEALLKRGCTATQVLHGAGALVDVASADEVSGVCAIVGALSALAAQKGTIY 123
 YY 121 LANKESTVTAAGQLMRTAENGATILPVDSHSAVFQALAGEDTA-----CVERVITTA 175
 D 124 LANKESTVTAAGQLMRTAENGATILPVDSHSAVFQALAGEDTA-----CVERVITTA 183
 YY 176 GGFPRDMSLERIRACTVAEQAHPNMSGORISIDSAMFNKALELLETREFFGFEPPDR 235
 D 184 GGFPLTADLNTFDSITPDQAVKHPNMRKGRKITSVDSATMANKGLELLEAHHLFCPPDKL 243
 YY 236 EAVVHPQSIYHAWVGFCDGGLMAHLGPADEMRHAIQFALNMPRGCEVPYARIDLAQIASLT 295
 D 244 EVVHPQSIYHAWVGFCDGGLMAHLGPADEMRHAIQFALNMPRGCEVPYARIDLAQIASLT 303
 YY 296 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 355
 D 304 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 363
 YY 356 VSTDPLFGKVPDALEEVLANMDHLARRAAEEAAG 388
 D 364 DFSDGI-----GDIGGLLAQDARTRAAQARATIG 391
 RESULT 12
 AAB01362
 ID AAB01362 standard; Protein; 473 AA.
 XX
 AC AAB01362;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE 1-deoxy-D-xylose-5-phosphate reductoisomerase sequence.
 XX
 KM 1-deoxy-D-xylose-5-phosphate reductoisomerase; transgenic plant;
 KM herbicide; isopentenyl diphosphate; mapping; breeding; phenotype;
 KM probe.
 XX
 OS Oryza sativa.
 XX
 PN WO200034448-A1.
 PD 15-JUN-2000.
 XX
 PF 03-DEC-1999; 99MO-US28616.
 XX
 PR 04-DEC-1998; 98US-0110865.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Lee J, Tao Y;
 XX
 DR WPI, 2000-431295/37.
 DR N-PSDB; AAA47405.
 XX
 PT Novel polynucleotide encoding isopentenyl diphosphate biosynthetic
 PT enzymes useful for producing transgenic plants with altered isopentenyl
 PT diphosphate levels and for selecting polynucleotides affecting

PT expression of the enzyme
 XX
 PS Claim 1; Page 41-43; 63pp; English.
 XX
 CC Nucleic acids encoding 1-deoxy-D-xylose-5-phosphate
 CC reductoisomerase are useful for creating transgenic plants with
 CC altered levels of plastid isopentenyl diphosphate in cells.
 CC The 1-deoxy-D-xylose-5-phosphate reductoisomerase polypeptides are
 CC useful as targets for identifying inhibitors of the enzyme which are
 CC useful as herbicides. Nucleic acid fragments of 1-deoxy-D-xylose-5-
 CC phosphate reductoisomerase are useful as probes for genetically and
 CC physically mapping genes useful in plant breeding for developing
 CC lines with desired phenotypes.
 XX
 SQ Sequence 473 AA;
 Query Match 37.4%; Score 742; DB 21; Length 473;
 Best Local Similarity 41.1%; Pred. No. 1.1e-66;
 Matches 164; Conservative 74; Mismatches 139; Indels 22; Gaps 7;
 YY
 2 RSLIFGATGSGESTFDLVMRKSGPEAFRTVALTGGRNIRLAEKARALKAEATVAHE 61
 D 77 KPIVIGSTGSGIGQTLDIVAE--NPDKFRVVALAAGSNVTLADQYKTFKPKLVAVRNE 134
 YY 62 DCLPALREALAGT--TEVAGGAQAIAEADR-ADMTSAIVGAAGLVGMRALKHGRT 118
 D 135 SLVDELKEALADCDMKPEIIPGEGVIEVARHPDAVTVVGVIGCAGLKEPTVAAIENGKD 194
 YY 119 LALNKESTVTAAGQLMRTAENGATILPVDSHSAVFQALAGEDTACVERVITTA 178
 D 195 IALNKESTVTAAGQLMRTAENGATILPVDSHSAVFQALAGEDTACVERVITTA 254
 YY 179 FRDMSLERIRACTVAEQAHPNMSGORISIDSAMFNKALELLETREFFGFEPPDR 238
 D 255 FRDMPVDELKEVVAADALKHPNMRKGRKITSVDSATLFNKGLIEVAHYLGAEDDIEIV 314
 YY 239 VHPQSIYHAWVGFCDGGLMAHLGPADEMRHAIQFALNMPRG--GEVVAARIDLAQIASLT 295
 D 315 IHPQSIYHAWVGFCDGGLMAHLGPADEMRHAIQFALNMPRG--GEVVAARIDLAQIASLT 374
 YY 296 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 355
 D 375 FKAPDNKRYPSMDLAIYAGRAGGTMTGVLNANKAVELFIDEKITGLDIFKVELTICA 434
 YY 356 ----VSTDPLFGKVPDALEEVLANMDHLARRAAEEAAGIR 390
 D 435 HRNELVTRP-----SLEEIHYDLWAR---EYAAASLQ 463
 RESULT 13
 AAY97197
 ID AAY97197 standard; Protein; 399 AA.
 XX
 AC AAY97197;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE A. thaliana partial 1-deoxy-D-xylose-5-phosphate reductoisomerase.
 XX
 KM 1-deoxy-D-xylose-5-phosphate reductoisomerase; peppermint; isoprenoid;
 KM synthesis; metabolism; chlorophyll; terpenoid; insecticidal; aroma;
 KM adhesive; ink; polymer; cytostatic; antiparasitic; plant.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200046346-A1.
 PD 10-AUG-2000.
 XX
 PF 27-JAN-2000; 2000MO-US02185.
 XX
 PR 03-FEB-1999; 99US-0118349.
 XX

PA (UNIM) UNIV WASHINGTON STATE RES FOUND.
 XX
 XX Croteau RB, Lange BM;
 XX
 XX WPI: 2000-524417/47.
 DR N-PSDB: AAA52152.
 XX
 XX Nucleic acid encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerase,
 PT useful for enhancing or altering isoprenoid synthesis and metabolism in
 PT plants
 XX
 XX Example 2: Page 54-55; 65pp; English.
 PS
 XX This is a partial 1-deoxy-D-xylulose-5-phosphate reductoisomerase
 CC isolated from an Arabidopsis thaliana flower bud cDNA library using a
 CC Mentha piperita (peppermint) probe (pMPDXRL, shown in AAA52151). The
 CC peppermint reductoisomerase can be used to transform plants to enhance
 CC or alter their isoprenoid synthesis and metabolism. The enzyme catalyzes
 CC the first committed step in the conversion of
 CC 1-deoxy-D-xylulose-5-phosphate to isopentenyl diphosphate, which in turn,
 CC is converted to a variety of molecules such as carotenoids, the prenyl
 CC side chains of chlorophyll, plastoquinone and tocopherols. The enzyme can
 CC therefore be used to enhance the production of chlorophyll, terpenoids,
 CC phytoalexins, toxins and detergent compounds to improve defence against
 CC pathogens, insects and other herbivores, enhance the production of
 CC monoterpene flavour and aroma profiles, prepare synthetic intermediates
 CC for industrial use such as the synthesis of adhesives, inks and polymers,
 CC improve the yield of natural pigments extracted from plants for medicinal
 CC or culinary uses, and to enhance the yield of compounds having
 CC anti-cancer properties such as vitamin A and vitamin E.
 CC
 XX
 XX Sequence 399 AA;
 SQ
 Query Match 37.4%; Score 741.5; DB 21; Length 399;
 Best Local Similarity 41.3%; Pred. No. 9.4e-67;
 Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;
 QY 2 RSLISFGATGSGESTFDLMRKGGPEAFRTVALTGRNIRRLAEMARALKAELAVTAHE 61
 Db 3 KRISIVSGTSGTGTLDIVAE--NPDKFRVVALAAGSNVTLADQVRRFKPALVAVRNE 60
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP-ADWMTSAIVGAAGLVPGMRALKHGT 118
 Db 61 SLINLEKEALADLDYKLEIIPGEOGVIEVARHPEAVTVVGTIVGAGLKPVAALAEAGKD 120
 QY 119 LALANKESIVTAGOLLMRTAENGATILPVDESHSAVFQALAGEPTACERYIITASGSP 178
 Db 121 IALANKETLIAGGPPVPLPLANKHNKILPDADESHSAIFQCIOGLEPAGARKIITILASGGA 180
 QY 179 FDMWLERIACCTVAEAOAHNPMWGMQORISIDSMFNKALELITREFEGEPRIEAV 238
 Db 181 FDMWVERKEKLVADALKHPPMNMNGKKTIVDSATLFNKGLEIVIAHNYLFGAEYDIEIV 240
 QY 239 VHPDSIVHAWGFCGGLMAHLGPRDMRAHIGFALNMPGR--GEVPAVIDLADIASLT 295
 Db 241 IHPDSIHSIMETQSSVLAQGLPDMRPLILYTWSPRPVPCSEVTWPRDLCKLGSLT 300
 QY 296 FOKPDEERPALRLARDVMAARGLSGAFAFNAAKETALDHFISGRIGFLDMAVVEETL-- 353
 Db 301 FKRPDVKKPPSMDLAVYAAGRAGGTMTGVLSAANEKAVEFIDEKISYLDIFRVELTCDK 360
 QY 354 --AGVSTDPLEFGKVPDALEEVILAMDLARRAA 383
 Db 361 HRNELVTSP-----SLEIVHYDLMAREYA 385
 RESULT 14
 AAB11412
 ID AAB11412 standard; Protein; 406 AA.
 XX
 AC AAB11412;
 XX
 DT 22-FEB-2001 (first entry)

XX
 DE A. thaliana DXPRI protein.
 XX
 KW DXPRI: 1-deoxy-D-xylulose-5-phosphate reductoisomerase; plant;
 KW tocopherol; Vitamin K; carotenoid; chlorophyll; polyterpene;
 KW 2-C-methyl-D-erythritol-4-phosphate; plastid isoprenoid.
 XX
 OS Arabidopsis thaliana.
 PN W0200065036-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 17-APR-2000; 2000WO-EP03465.
 XX
 PR 27-APR-1999; 99DE-1018949.
 PA (BADI) BASF AG.
 PI Lichtenhaler H, Schwender J, Reindl A, Herbers K;
 DR WPI: 2000-687330/67.
 DR N-PSDB: AAC81871.
 XX
 PT Nucleic acids encoding plant 1-deoxy-D-xylulose-5-phosphate
 PT reductoisomerase, useful for producing transgenic plants with
 PT increased content of e.g. tocopherols and carotenoids -
 PS disclosure: Page 35-37; 41pp; German.
 XX
 PS Sequence 406 AA;
 SQ
 Query Match 37.4%; Score 741.5; DB 21; Length 406;
 Best Local Similarity 41.3%; Pred. No. 9.6e-67;
 Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;
 QY 2 RSLISFGATGSGESTFDLMRKGGPEAFRTVALTGRNIRRLAEMARALKAELAVTAHE 61
 Db 10 KRISIVSGTSGTGTLDIVAE--NPDKFRVVALAAGSNVTLADQVRRFKPALVAVRNE 67
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP-ADWMTSAIVGAAGLVPGMRALKHGT 118
 Db 68 SLINLEKEALADLDYKLEIIPGEOGVIEVARHPEAVTVVGTIVGAGLKPVAALAEAGKD 127
 QY 119 LALANKESIVTAGOLLMRTAENGATILPVDESHSAVFQALAGEPTACERYIITASGSP 178
 Db 128 IALANKETLIAGGPPVPLPLANKHNKILPDADESHSAIFQCIOGLEPAGARKIITILASGGA 187
 QY 179 FDMWLERIACCTVAEAOAHNPMWGMQORISIDSMFNKALELITREFEGEPRIEAV 238
 Db 188 FDMWVERKEKLVADALKHPPMNMNGKKTIVDSATLFNKGLEIVIAHNYLFGAEYDIEIV 247
 QY 239 VHPDSIVHAWGFCGGLMAHLGPRDMRAHIGFALNMPGR--GEVPAVIDLADIASLT 295
 Db 248 IHPDSIHSIMETQSSVLAQGLPDMRPLILYTWSPRPVPCSEVTWPRDLCKLGSLT 307
 QY 296 FOKPDEERPALRLARDVMAARGLSGAFAFNAAKETALDHFISGRIGFLDMAVVEETL-- 353
 Db 308 FKRPDVKKPPSMDLAVYAAGRAGGTMTGVLSAANEKAVEFIDEKISYLDIFRVELTCDK 367
 QY 354 --AGVSTDPLEFGKVPDALEEVILAMDLARRAA 383

DB	368	HNRELVTSP-----SLEEIVHYDMAREYA	392
RESULT 15			
AAB01370			
ID	AAB01370	standard; Protein; 406 AA.	
XX			
AC	AAB01370;		
XX			
DT	16-NOV-2000	(first entry)	
XX			
DE	1-deoxy-D-xylulose-5-phosphate reductoisomerase.		
XX			
KN	1-deoxy-D-xylulose-5-phosphate reductoisomerase; transgenic plant;		
KN	herbicide; isopentenyl diphosphate; mapping; breeding; phenotype;		
KN	probe.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	MO200034448-AL.		
XX			
PD	15-JUN-2000.		
XX			
PF	03-DEC-1999;	99WO-US28616.	
XX			
PR	04-DEC-1998;	98US-0110865.	
XX			
PA	(DDPO) DU PONT DE NEMOURS & CO E. I.		
XX			
PI	Cahoon RE, Lee J, Tao Y;		
XX			
DR	WPI, 2000-431295/37.		
XX			
PI	Novel polynucleotide encoding isopentenyl diphosphate biosynthetic		
PR	enzymes useful for producing transgenic plants with altered isopentenyl		
PT	diphosphate levels and for selecting polynucleotides affecting		
PT	expression of the enzyme		
XX			
PS	Example 3; Page 57-58; 63pp: English.		
XX			
CC	Nucleic acids encoding 1-deoxy-D-xylulose-5-phosphate		
CC	reductoisomerase are useful for creating transgenic plants with		
CC	altered levels of plastid isopentenyl diphosphate in cells.		
CC	The 1-deoxy-D-xylulose 5-phosphate reductoisomerase polypeptides are		
CC	useful as targets for identifying inhibitors of the enzyme which are		
CC	useful as herbicides; Nucleic acid fragments of 1-deoxy-D-xylulose-5-		
CC	phosphate reductoisomerase are useful as probes for genetically and		
CC	physically mapping genes useful in plant breeding for developing		
CC	lines with desired phenotypes.		
XX			
SO	Sequence 406 AA;		
Query Match	37.4%; Score 741.5; DB 21; Length 406;		
Best Local Similarity	41.3%; Pred. No. 9.6e-67;		
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;			
QY	2 RSLSTFGATGSGESTFIDLWMKGGPEAFRYALTGGRNIRLAEMARLKAELAVTAHE 61		
DB	10 KRISTVGTSGTSGITDILIAE--NPKDFRYVALAAGSVTLADQVRRFKPALVAVRE 67		
QY	62 DCLPLRLAAGTG--TEVAGAAQALAEADNP-ADWTMSALVGAAGLVPGRAKLGHT 118		
DB	68 SLINELKALALADLYKLEIRPEQGVIEAARRPEVTVVGTGAGCAKPTVAAALEAGD 127		
QY	119 LALAKESLVTAGOLIMLTAEONGATILPVDSHSAVFOALGEPDTACVERVIITASGP 178		
DB	128 IALANKETLLINGRPVYLPLANKHNKILIPADSEHSATFOCIDGLPEGALRKITILIASGGA 187		
QY	179 FRDMSLERACTVAEAQAHNPMWSGQRISIDSASMPKFALELITREFGEPPRIEAV 238		
DB	188 FRDMVEVETLKEVYADALAKHPNMWNGKKITVDSATLFLFKGLVEIVAHNLFAGAEYDIEIV 247		
QY	239 VHPSIVAMVGFCDGGLMAHLGRPMADHNAIGFALNMFGR--GEVPAARIDLAOIASIT* 295		

```

Db      248  IHPSIIHSMLETTDSSVLAIGMPDMRLPLITYMSNPDRKPCSEYVWPPRLDCKLGSLT 307
Oy      296  FQKPDDEEFPALRLARDYMAARGLSGAAFNAKEIALDHPITAGRIGFLDMAAAVEETL-- 353
Db      308  FKKPDNVVYPSMDLAIYAAGRAGGTMTGVLSTANAEKAVEMFIDEKISYLDIFKVELTCDK 367
Oy      354  --AGVSTDPLFGKVPDALEEVLDHDLARRAA 383
Db      368  HKNELVTSF-----SLEELVHYHDLNAREYA 392

```

Search completed: April 15, 2003, 13:57:45
Job time : 39 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:29 ; Search time 21 Seconds

(without alignments)
1803.665 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRSLSIFGATGSGESTFDL.....MDHLARRAEAAAGLRQKR 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	994	50.1	397	2 E97672	1-deoxy-D-xylose
2	994	50.1	397	2 AC2897	1-deoxy-D-xylose
3	899.5	45.3	399	2 H87486	1-deoxy-D-xylose
4	775.5	39.1	396	2 E83188	1-deoxy-D-xylose
5	750.5	37.8	394	2 A81229	1-deoxy-D-xylose
6	743.5	37.5	394	2 B82000	1-deoxy-D-xylose
7	741.5	37.4	406	2 T52570	1-deoxy-D-xylose
8	740	37.3	396	2 H82728	1-deoxy-D-xylose
9	733	36.9	398	2 G90650	1-deoxy-D-xylose
10	733	36.9	398	2 G85501	1-deoxy-D-xylose
11	732	36.9	398	2 E64741	1-deoxy-D-xylose
12	725	36.5	398	2 AG0128	1-deoxy-D-xylose
13	717	36.1	397	2 A64014	1-deoxy-D-xylose
14	708	35.7	398	2 AF0529	1-deoxy-D-xylose
15	707	35.6	399	2 AH2349	1-deoxy-D-xylose
16	702.5	35.4	380	2 AE1239	1-deoxy-D-xylose
17	695.5	35.0	394	2 S76331	1-deoxy-D-xylose
18	693.5	35.0	385	2 E97121	1-deoxy-D-xylose
19	682	34.4	402	2 D82099	1-deoxy-D-xylose
20	673	33.9	380	2 A11601	1-deoxy-D-xylose
21	670	33.8	388	2 B69881	1-deoxy-D-xylose
22	640	32.3	396	2 F75388	1-deoxy-D-xylose
23	631	31.8	393	2 C71304	1-deoxy-D-xylose
24	628	31.7	436	2 A70923	1-deoxy-D-xylose
25	617	31.1	398	2 F84957	1-deoxy-D-xylose
26	611.5	30.8	379	2 G81122	1-deoxy-D-xylose
27	608.5	30.7	379	2 A71562	1-deoxy-D-xylose
28	598.5	30.2	379	2 C81578	1-deoxy-D-xylose
29	596.5	30.1	379	2 G86533	1-deoxy-D-xylose

30	596.5	30.1	379	2 B72091	ct071 hypothetical
31	595.5	30.0	365	2 E83952	1-deoxy-d-xylose
32	588	29.6	406	2 A87107	hypothetical prote
33	565	28.5	380	2 F70336	conserved hypotet
34	550.5	27.7	376	2 B72321	conserved hypotet
35	488	24.6	368	2 H64546	conserved hypotet
36	477	24.0	368	2 G71961	hypothetical prote
37	468	23.6	356	2 G81278	probable 1-deoxy-D
38	133.5	6.7	412	2 B87586	hypothetical prote
39	125.5	6.3	572	2 H75271	conserved hypotet
40	123.5	6.2	1731	2 B98241	hypothetical prote
41	123.5	6.2	1731	2 AB3045	ice nucleation pro
42	119	6.0	538	2 B70877	hypothetical prote
43	117	5.9	520	2 E87621	hypothetical prote
44	116	5.8	547	2 H83018	dihydroliopamide a
45	115.5	5.8	434	2 AE3550	homoserine dehydro

ALIGNMENTS

RESULT 1

E97672

1-deoxy-D-xylose 5-phosphate reductoisomerase (dxp reductoisomerase) [imported] - A

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97672

R:Gooner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm

A.: Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Martels,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E97672

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <RUR>

A:Cross-references: GB:AE007869; PIDN:AAK88334.1; PID:jl5157812; GSPDB:GND0169

C:Genetics:

A:Gene: AGR C.4736

A:Map position: circular chromosome

C:Superfamily: conserved hypothetical protein HI0807

Query Match	Best Local Similarity	Score	DB 2;	Length	397;
Matches	197;	Conservative	64;	Mismatches	126;
Indels	4;	Gaps	1;		
QY	2	RSLSIFATGSGISGFTFDLVMKRGPEAFRTVALTGRNTRLAEMARALKAEIATVAHE	61		
DB	9	RKLITLIGSTGISTNTLDVVRQLOGRGDFEIMALTGAGNIALLAEOARREGAQLATVADD	68		
QY	62	DCUPALREALAGTGEVAGGAQAIAEAADRPADMTGAIYGAAGLVGMRALKHGTAL	121		
DB	69	DKYALKSALAGTGIKRAAGAGLEEAASMDAGVMAAIAIGTGLATVTRAAKAGADIAL	128		
QY	122	ANKESIVTAQQLMRTAENGATILPVDSHSAVFOALAGEDNACVETVITASGGEFRT	181		
DB	129	ANKKCLVAGDVFLRTVKOGGGRILIPVDSHSAIFQCLTGEYKQAVRIVLTASGGEFRT	188		
QY	182	WLELRITACTYVAEQAPNMSMGORISIDASMSKALELIERREFGFEPRDREAVVHP	241		
DB	189	WSRDEMSNVTAADIAFARHPNMSMGLKVSIGSASMEKLEMEIAYLFDLPDQDVIVHP	248		
QY	242	QSTYHAWVGFCGDMHILPADMRHIGFALNMPGGEVPAVRIIDLAQIASLTFOKPDE	301		
DB	249	QSTIHSWVGITDGSYIAQLQSPDKRTAISTALYPERGNLSVERLDPAKLARDFEAPDE	308		
QY	302	EREPALRLADYVMAKSLGAFAAFAEIALDHFIAIRIGFLDMAAAVEETLAGVSTPDL	361		
DB	309	ARFALALAMALERGLOGAALNAAEFTAFHAFVAGIGTFLDMAELVEYVMDRMHD---	365		
QY	362	FGKYPDALEEVYLANDHLARRAAEAGLRQO	392		
DB	366	GRTAETMDVFSADDEARRHALLLTIATKER	395		

A:Molecule type: mRNA
A:Residues: 1-406 <SCH>
A:Cross-references: EMBL:AJ242588; NID:94886306; PIDN:CAB4344.1; PID:94886307
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: dxt
C:Superfamily: conserved hypothetical protein H10807

Query Match 37.4%; Score 741.5; DB 2; Length 406;
Best Local Similarity 41.3%; Pred. No. 1.4e-45;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLSTFGATGSGESTFDLVNKRKGPEAFRTVALTGGRNIRRLAEMARALAEAVTAH 61
D 10 KRISVGTSGISGTDLDIVAE--NPDKFRVVALAGSNVLLADQVRRFPALVAANE 67
QY 62 DCLPALREALGTG--FTVAGGAOAIARADRP--ADWTMSAIVGAGVPGPRAKHKRT 118
D 68 SLINELKALADLDYKKEIIPGEGVIEVARHPEAVYVGTIGVCGAKLPVAAIEAGKD 127
QY 119 LALANKESLVYAGQLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITASGP 178
D 128 IALANKETLIGAGFVLEPLANKHNVKILPADSEHSAIFQCIQGLPEGALRKILITASGA 187
QY 179 FRDMSLERIRACTVAEAOAHNPMMSGORISIDSASMEKALELITREFFGPEPRIEAV 238
D 188 FRDPVPEVLEKVKYADALKHNNMNGKKITVDSATLFFKGLVEIAHYLEFGAEYDIDIV 247
QY 239 VHPDSIVHNAVGFCDGGLMAHLGPADMRHAIIGFALNMPGR--GEVPARIDLAQIASLT 295
D 248 IHPSIISHMETQDSSVLAQGLMDMLPLITMSPDRPCSEVTMPRIDLCLSLT 307
QY 296 FQKDEEFPALRLARDVMAARGLSGAAPNAKEIALDHFITAGRTGLDMAAVEETL-- 353
D 308 FKRPDMVKYPMMDIAYAGRAGGTWTGVLASANEKAVMEFIDEKISTYDIKVVETLTDCK 367
QY 354 --AGVSTDPLEFGVYDPALEEVTLAMDHLARRAA 383
D 368 HRNELVTS-----SLEIVHYDLMAREYA 392

RESULT 8
H82728
1-deoxy-D-xylose 5-phosphate reductoisomerase Xf1048 [imported] - Xylella fastidiosa
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <SIM>
A:Cross-references: GB:AE003942; GB:AE003849; NID:99105990; PIDN:AAF83858.1; GSPDB:GN001
A:Experimental source: strain 945c
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brites, M.R.S.; Bueno, M.R.P.; Camargo, L.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Melo, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitzima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
Chado, J.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.V.; Savasak
Rodrigues, V.; Rosa, A.C.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sattelli, R.V.; Savasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshano, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:

A:Gene: Xf1048
C:Superfamily: conserved hypothetical protein H10807

Query Match 37.3%; Score 740; DB 2; Length 396;
Best Local Similarity 42.4%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 154; Indels 10; Gaps 5;

QY 1 MRSLSIFGATGSGESTFDLVNKRKGPEAFRTVALTGGRNIRRLAEMARALAEAVTAH 60
D 5 IRNVAIVGATGSGAALDVLARH--PROFHVSLLAQGRVALLALCHTRPDHVAICD 62
QY 61 EDCPLALREAL--AGTGVAGGAOAIAR--AADRADMTSAIYAGAGVGMRLKGR 117
D 63 ATLYTLLDGGNAGLAKAVAGENAELVASTCDDVVAIYAGALHSTLAAARAGK 122
QY 118 TLALANKESLVYAGQLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITASGP 177
D 123 RLLANKESLVLAGMLMRKESISGAELIIPIDSEHNAIFQCLRSRTTGGVIRITLTAAG 182
QY 178 PFRDMSLERIRACTVAEAOAHNPMMSGORISIDSASMEKALELITREFFGPEPRIEA 237
D 183 PFRGNHRTMLAKITPTQAMAHPTWMSGPKISVDSATLNNKGLVEIAHHLFGLPSEQIDV 242
QY 238 VHPDSIVHNAVGFCDGGLMAHLGPADMRHAIIGFALNMPGGEVPAVIDLAQIASLTFO 297
D 243 LVHPSLVHSLVEIFDGSSTLQSLPDMRTTLVAGLSMPERISGCVPGDLMKHRLDFE 302
QY 298 RPDSEFPALRLARDVMAARGLSGAAPNAKEIALDHFITAGRTGLDMAAVEETLACVS 357
D 303 RPDTEFSCRLARAMOTGTAPAVLNAANEIASATLQGRIGRTLPALIEHNL--T 359
QY 358 TDPLFGKVPDALEEVTLAMDHLARRAAEEA 386
D 360 TLPRY--EADTLETLTVDTERTRTTHAA 386

RESULT 9
G90650
1-deoxy-D-xylose 5-phosphate reductoisomerase [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasekara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90650
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833598.1; PID:913359631; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ecs0175
C:Superfamily: conserved hypothetical protein H10807

Query Match 36.9%; Score 733; DB 2; Length 398;
Best Local Similarity 40.7%; Pred. No. 5.5e-45;
Matches 162; Conservative 72; Mismatches 144; Indels 20; Gaps 5;

QY 1 MRSLSIFGATGSGESTFDLVNKRKGPEAFRTVALTGGRNIRRLAEMARALAEAVTAH 60
D 1 MKQTLITGSGISGTDLDV--RHNPFRVVALVAGKNTVAVGCLSPRYAVHDD 58
QY 61 EDCPLALREAL--AGTGVAGGAOAIAR--AADRADMTSAIYAGAGVGMRLKGR 117
D 59 EASKLKLTKMLQOGCSRTVEVSGQACDMALAEVDVQVMAIYAGAGLPTLAINIRGK 118
QY 118 TLALANKESLVYAGQLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITACV 167
D 119 TILANKESLVTCGLRFDPAVKQSKAQLLPDSEHNAIFQSLPPQICNNLGYADLENGV 178

C:Superfamily: conserved hypothetical protein HI0807
C:Keywords: oxidoreductase

Query Match 36.5%; Score 725; DB 2; Length 398;
Best Local Similarity 40.7%; Pred. No. 2,1e-44;
Matches 162; Conservative 69; Mismatches 147; Indels 20; Gaps 5;

QY 1 MRSLSIFGATSGISTEFTDLVVRKGGPEAFRTVALTGGRNIRLAEAKALAEAYTAH 60
D 1 MKQLTILGSTGISTSLVY--RANPELFKVTALVAGNRVEMAOQLERSPRYAASD 58
QY 61 EDCPLALREALA--GTGEVAGGAQALAE--AADRPADMTSAIYGAAGLVGMRALKGR 117
D 59 EHSKSKSLRLAEQSGDTEVSGETAACELALDDVDQMAIYAGLIPSTLAIRAGK 118
QY 118 TLALANKESLYTAGQLMRTAQENGATILPVDESHSAVFOAL-----AGEDTACV 167
D 119 QVLANKESLITGCKLFPDEVKRSRAQLPIDSERHAIFOGLPERIQOLGYSLSLNGV 178
QY 168 ERVITTAGSGPFRDMSLERIACVTAEAQAHPNMSMGORISIDSASFNKALELITREF 227
D 179 SRILTGSGGPFRETPLSQFSDVTPDQCAHPNMSMGRIKISVDSATMKNKGLLEYEARWL 238
QY 228 FGFEPRLEAVVHPOSIYHNAVVGCDGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 287
D 239 FNASAEQLEVLHPQSVHSWRYHDGSLQMGTPDMRTTIAHAMAYPMRVSSGAPLD 298
QY 288 LAQIASLTFQKPDDEERFPALRLARDVMAARGLSGAFAFAAKEIALDHFIAGRIGFLDMAA 347
D 299 FCKVAGLTFTPDYQRYPCKLALDAGNAGQAATTALNAANEISVMAFLDSKIRTDIEV 358
QY 348 VVEETLAGVSTDPLEFGKVPDALEVLAMDHLARRAAE 385
D 359 INRTVEG---LLSEPTVSEEVLYIDRRARVAAQ 391

RESULT 13

A64014
conserved hypothetical protein HI0807 - Haemophilus influenzae (strain Rd KW20).

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence, revision 18-Aug-1995 #text_change 29-Sep-1999

C:Accession: A64014

R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Funtmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: A64014
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-397 <TRIG>

A:Cross-references: GB:U32763; GB:I42023; NID:g1573817; PIDN:AC22466.1; PID:g1573819; T C:Superfamily: conserved hypothetical protein HI0807

Query Match 36.1%; Score 717; DB 2; Length 397;
Best Local Similarity 40.2%; Pred. No. 7.7e-44;

Matches 159; Conservative 72; Mismatches 145; Indels 20; Gaps 5;

QY 2 RSLISFGATSGISTEFTDLVVRKGGPEAFRTVALTGGRNIRLAEAKALAEAYTAH 61
D 4 QNITLILGSTGISTSLVY--ENNPQRYHAFALVGGKNAEMEQCIKFRPHFAALDDV 61
QY 62 DCLPALREALAG--GTGEVAGGAQALAEADRP--ADMTSAIYGAAGLVGMRALKHGR 118
D 62 NAAKILKEKLAHRIETEVLAGRAICELAHAPDADIMASIVGAGLIPLSAVKAGKR 121
QY 119 LALANKESLYTAGQLMRTAQENGATILPVDESHSAVFOALAGE-----DTACVE 168
D 122 VLLANKESLYTAGQLMRTAQENGATILPVDESHSAVFOALAGE-----DTACVE 168
QY 169 RVITTAGSGPFRDMSLERIACVTAEAQAHPNMSMGORISIDSASFNKALELITREF 228

D 182 KIILTGSGGPFRRYPLPEQFTNITPEQAAVHAPNMSMGRIKISVDSATMKNKGLLEYEARWL 241
QY 229 GFEPRLEAVVHPOSIYHNAVVGCDGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 288
D 242 NASAEQLEVLHPQSVHSWRYHDGSLQMGTPDMRTTIAHAMAYPMRVSSGAPLD 298
QY 289 LAQIASLTFQKPDDEERFPALRLARDVMAARGLSGAFAFAAKEIALDHFIAGRIGFLDMAA 348
D 302 FKIELFTIEDPNFNPULKIALDAFAAGQYATTAMNANEIYAOAFLDRIGFMDIAKI 361
QY 349 VVEETLAGVSTDPLEFGKVPDALEVLAMDHLARRAAE 384
D 362 NSKTIERTISPYTI-----QNIIDVLEIDAQAREIAK 392

RESULT 14

AF0529

1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.-) [imported] - Salmonella

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence, revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AF0529

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Mout, S.; O'Gaora, P. Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AF0529
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08678.1; PID:g16501501; GSPDB:GN00176

C:Genetics:

A:Gene: dxr

C:Superfamily: conserved hypothetical protein HI0807

C:Keywords: oxidoreductase

Query Match 35.7%; Score 708; DB 2; Length 398;
Best Local Similarity 40.2%; Pred. No. 3.4e-43;
Matches 160; Conservative 72; Mismatches 146; Indels 20; Gaps 6;

QY 1 MRSLSIFGATSGISTEFTDLVVRKGGPEAFRTVALTGGRNIRLAEAKALAEAYTAH 60
D 1 MKQLTILGSTGISTSLVY--NPDSEFVIALVAGKNVARMADQCLERSPRYAASD 58
QY 61 EDCPLALREALA--GTGEVAGGAQALAE--AADRPADMTSAIYGAAGLVGMRALKHGR 117
D 59 TSSAEQLEVLHPQSVHSWRYHDGSLQMGTPDMRTTIAHAMAYPMRVSSGAPLD 298
QY 118 TLALANKESLYTAGQLMRTAQENGATILPVDESHSAVFOAL-----AGEDTACV 167
D 119 TILANKESLYTAGQLMRTAQENGATILPVDESHSAVFOAL-----AGEDTACV 167
QY 168 ERVITTAGSGPFRDMSLERIACVTAEAQAHPNMSMGORISIDSASFNKALELITREF 227
D 179 TSILTGSGGPFRETPLMCDLAAATPDQCRHPNMSMGRIKISVDSATMKNKGLLEYEARWL 238
QY 228 FGFEPRLEAVVHPOSIYHNAVVGCDGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 287
D 239 FNASAEQLEVLHPQSVHSWRYHDGSLQMGTPDMRTTIAHAMAYPMRVSSGAPLD 298
QY 288 LAQIASLTFQKPDDEERFPALRLARDVMAARGLSGAFAFAAKEIALDHFIAGRIGFLDMAA 347
D 299 FCKLSALTFSPADYQRYPCKLALDAGNAGQAATTALNAANEIYAOAFLDRIGFMDIAKI 361
QY 348 VVEETLAGVSTDPLEFGKVPDALEVLAMDHLARRAAE 385
D 359 L---NLAVLERMD--HEPASVDVLOYDALAREVARK 391

RESULT 15
AH2349
deoxyxylulose 5-phosphate reductoisomerase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH2349
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriuchika, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MIMD:21595285; PMID:11759840
A:Accession: AH2349
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <KUR>
A:Cross-References: GB:BA000019; PIRN:BA06050.1; PID:g17133487; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4351
C:Superfamily: conserved hypothetical protein H10807

Query Match 35.6%; Score 707; DB 2; Length 399;
Best Local Similarity 38.0%; Pred. No. 4e-43;
Matches 146; Conservative 88; Mismatches 142; Indels 8; Gaps 4;

QY 1 MRSLSTFGATGSGTSGTSTGPDLMVRKGPPEAFRTVALTGRNIRRLAEARALKAEIAVTAH 60
DB 2 VKSITIVGSGTSGTGTLDIVSOY--PDQRIYGLAGSNVEMIAEQIRFROIATISA 59
QY 61 EDLPALREALAGTGE--VAGGAQAIAEAAD--RPADWTMSATVGAAGLVPGRALKHGR 117
DB 60 AEKLPALQAAIKDLPQPIILGGEAGVIEVARGDATVTVTGCGAGLPTIAIEAGK 119
QY 118 TLALAKESIVTGAQLMRTAENGATILPVDSEHSNVAFOALAGEDTACVETVITRASG 177
DB 120 DIALANKETLIAGGPVPLPLVEKHGVKLLPADSEHSAIFQCIQGVPRKGLKILLTASGG 179
QY 178 PFRDMSLERIRACTVAEQAHPNWSMGORISDSAMFNKALELIEETREFGFEPPRIEA 237
DB 180 AFRDMDVERIAETVSDALKHPNWSMGKRTTYSATILMNKGLEIVENHFLFGLDYQDIET 239
QY 238 VVHPOSIVHNAVGFCDGGLMAHLGPADMRHAIQFALNMPGREGVPAVARIDLAQIASLTFO 297
DB 240 VIHPOSIIHSLIELQDTSVLAQLGMPDMRLPLLYALSMPEIRIYTDWERNLVKAGNLTFR 299
QY 298 KPDEERFPALRLARDVMAARGLSGAIFNAKETALDHFITAGRICFLDMAVVEETLAGVS 357
DB 300 EPDHQKPCMQQLAYAAAGRAGSGMPAVLNANAEQVVALFDEKIKFLDIPRCIELVCDRHQ 359
QY 358 TDPLFGKVPDALEVLAMHLARR 381
DB 360 ND---NCANPSLDDIILAADQMARQ 380

Search completed: April 15, 2003, 13:58:13
Job time : 23 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:25 ; Search time 25 Seconds

(without alignments)
653.667 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRLSIFGATGSGESTFDL.....MDHLARRAAEAGGRQKR 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1024	51.6	391	1 DXR_RHIME	0921p6 rhizobium m
2	984	50.1	397	1 DXR_AGR75	08uc86 agrobacteri
3	899.5	45.3	399	1 DXR_CACR	09a709 caulobacter
4	875	44.1	388	1 DXR_ZYMO	09x7f2 zymomonas m
5	797.5	40.2	393	1 DXR_PALSO	08x2f2 ralsstonia s
6	775.5	39.1	396	1 DXR_PSEAE	09x2u6 pseudomonas
7	750.5	37.8	394	1 DXR_NEITB	09j1g8 neisseria m
8	743.5	37.5	394	1 DXR_NEITB	09j1g3 neisseria m
9	741.5	37.4	477	1 DXR_NEITB	09x7s9 arabidopsis
10	740	37.3	396	1 DXR_XYLF	09x210 xyella fas
11	733	36.9	398	1 DXR_EC057	08x8y1 escherichia
12	732	36.9	398	1 DXR_EC057	08x8y1 escherichia
13	725	36.5	398	1 DXR_YERPE	08x8y1 escherichia
14	717	36.1	397	1 DXR_HAEN	08x8y1 escherichia
15	716	36.1	405	1 DXR_PASW	08x8y1 escherichia
16	715	36.0	398	1 DXR_SALT	08x8y1 escherichia
17	708	35.7	398	1 DXR_SALT	08x8y1 escherichia
18	707	35.6	399	1 DXR_ANASP	08x8y1 escherichia
19	702.5	35.4	380	1 DXR_LISMO	08x8y1 escherichia
20	695.5	35.1	394	1 DXR_SYNY	08x8y1 escherichia
21	693.5	35.0	385	1 DXR_CIOAB	08x8y1 escherichia
22	685	34.5	402	1 DXR_SYNE	08x8y1 escherichia
23	684.5	34.5	402	1 DXR_SYNE	08x8y1 escherichia
24	683.5	34.5	401	1 DXR_SPCOC	08x8y1 escherichia
25	682	34.4	402	1 DXR_VIBCH	08x8y1 escherichia
26	673	33.9	380	1 DXR_LISIN	08x8y1 escherichia
27	671.5	33.8	384	1 DXR_CLOPE	08x8y1 escherichia
28	670	33.8	388	1 DXR_BACSU	08x8y1 escherichia
29	654.5	33.0	475	1 DXR_MENPI	08x8y1 escherichia
30	643.5	32.4	382	1 DXR_BACHD	08x8y1 escherichia
31	640	32.3	396	1 DXR_DEIRA	08x8y1 escherichia
32	631	31.8	376	1 DXR_TREPA	08x8y1 escherichia
33	628	31.7	413	1 DXR_MYCTU	08x8y1 escherichia

34	617	31.1	398	1 DXR_BUCAI	P57329 buchnera ap
35	611.5	30.8	379	1 DXR_CHLMU	09x7u8 chlamydia m
36	608.5	30.7	379	1 DXR_CHLR	084074 chlamydia t
37	596.5	30.1	379	1 DXR_CHLR	09x7u8 chlamydia p
38	588	29.6	406	1 DXR_MYCLE	09x7u8 chlamydia p
39	568	28.6	390	1 DXR_FUSNN	09x7u8 chlamydia p
40	565	28.5	380	1 DXR_AQUAE	08x722 fusobacteri
41	550.5	27.7	376	1 DXR_THEMA	09x722 fusobacteri
42	488	24.6	368	1 DXR_HELPY	09x722 fusobacteri
43	477	23.6	356	1 DXR_HELPY	09x722 fusobacteri
44	468	23.6	356	1 DXR_HELPY	09x722 fusobacteri
45	140.5	7.1	2564	1 SPOC_HUMAN	09x722 fusobacteri

ALIGNMENTS

RESULT 1
ID DXR_RHIME STANDARD; PRT; 391 AA.
AC 0921p6;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
GN DXP OR R02988 OR SMC03105.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Gallbert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP) (by similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
-> 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AL591792; CAC47567.1;
CC InterPro: IPR003821; DXP_reductoisomase.
CC Pfam: PF02670; DXP_reductoisom; 1.
CC TIGRPFAM: TIGR00243; Dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 14 21 NADPH (POTENTIAL).
SQ SEQUENCE 391 AA; 41741 MW; 28036C08F026757D CRC64;
Query Match 51.6%; Score 1024; DB 1; Length 391;
Best Local Similarity 53.9%; Pred. No. 4.2e-65;
Matches 207; Conservative 60; Mismatches 113; Indels 4; Gaps 1;
QY 2 RSLISFGATGSGESTFDLVMRKGPGEAFRTVALTGGNRIRLAEMARALKAELAVTAHE 61

```

Db 9 RRLTILSTGSGTSTLVDIERLGRDRFETALTGNNINILAEQARRIGAEIATVAD 68
QY 62 DCLPALREALGCTGEVAGGAOATAEADRPADMTMSAIVGACIVPGRMLKRGRTAL 121
Db 69 DRYELKDALSGSGTGEVAGSGLEAEARDGAWMAIVGAGPPLAARRKADIAL 128
QY 122 ANKESLVYAGOLMRTAENGATILPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 181
Db 129 ANKCVISAGSLFIDVAEAGGRLPVDSEHNAIFOVLENGORRAVEIVITLASSGPPRT 188
QY 182 WSLERIRACTYAEAOAHPNMSMGORISIDSAMFNKALELIEETBEFGFEPDRILEAVVHP 241
Db 189 KTLDEMVRVTDVARRAHPNMSMGKISIDSAMFNKALEMIEARHLFRLEPQIEVIVHP 248
QY 242 OSIYHAWGFCDDGLMAHLGPAIDMRHAIGFALNMPGREGVAVARIDLAQISLTFORPDE 301
Db 249 OSIYHAWGTYDGSYVLAQOLGCPDMRTAIGALSTPKRCDLVERLDFAIRLDEEAPDE 308
QY 302 ERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAAAVEETLAGVSTPDL 361
Db 309 VRPAIKLARAMEGGVQGAVALNGAKRTALDAFIKGRIGFLAAMAEIVEXKMDGLAGLP- 367
QY 362 FGKVPDALEEVLANDHARAAAE 365
Db 368 ---AATSMVDVFAADERARRAAAE 388

```

RESULT 2

```

DXR_AGR5 ID DXR_AGR5 STANDARD; PRT; 397 AA.
AC Q98C86;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
  reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR ATU2612 OR AGR_C_4736.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
  Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; Pubmed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-Y., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
  C58."
RT Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; Pubmed=11743194; Miller N., Blanchard M.,
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollan C., Allinger M., Doughy D., Scott C., Iappas C., Markelz B.,
RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
  Agrobacterium tumefaciens C58."
RT Science 294:2323-2328(2001).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
  of 1-deoxy-D-xyulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
  4-phosphate (MEP) (By similarity).

```

```

CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nomenclonate terpenoid biosynthesis pathway; second
  step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE009208; AAL43593.1; -.
KW EMBL: AE008174; AAK8334.1; -.
FT NP_BIND 14 21 NADPH (POTENTIAL).
SQ SEQUENCE 397 AA; 42303 MW; 0A24C408C2D47094 CRC64;

```

Query Match 50.1%; Score 994; DB 1; Length 397;

Best Local Similarity 50.4%; Pred. No. 5; Se-63;

Matches 197; Conservative 64; Mismatches 126; Indels 4; Gaps 1;

```

QY 2 RSLIFGATGSGTSTLVDVNRKCGPEAFRTVALTGRNIRRLAEMARALAEIATVARE 61
Db 9 RRLTILSTGSGTSTLVDVNRKCGPEAFRTVALTGRNIRRLAEMARALAEIATVARE 68
QY 62 DCLPALREALGCTGEVAGGAOATAEADRPADMTMSAIVGACIVPGRMLKRGRTAL 121
Db 69 DRYELKDALSGSGTGEVAGSGLEAEARDGAWMAIVGAGPPLAARRKADIAL 128
QY 122 ANKESLVYAGOLMRTAENGATILPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 181
Db 129 ANKCVISAGSLFIDVAEAGGRLPVDSEHNAIFOVLENGORRAVEIVITLASSGPPRT 188
QY 182 WSLERIRACTYAEAOAHPNMSMGORISIDSAMFNKALELIEETBEFGFEPDRILEAVVHP 241
Db 189 WSRPESVNTADIRAHPNMSMGKISIDSAMFNKALEMIEARHLFRLEPQIEVIVHP 248
QY 242 OSIYHAWGFCDDGLMAHLGPAIDMRHAIGFALNMPGREGVAVARIDLAQISLTFORPDE 301
Db 249 OSIYHAWGTYDGSYVLAQOLGCPDMRTAIGALSTPKRCDLVERLDFAIRLDEEAPDE 308
QY 302 ERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAAAVEETLAGVSTPDL 361
Db 309 ARPALRLARALEGGVQGAVALNGAKRTALDAFIKGRIGFLAAMAEIVETWDRMHD--- 365
QY 362 FGKVPDALEEVLANDHARAAAEALGRQ 392
Db 366 -GRTAETMDVFSADDEARRRRLAELIATKEK 395

```

RESULT 3

```

DXR_CAUCR ID DXR_CAUCR STANDARD; PRT; 399 AA.
AC Q9A709;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
  reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR CCI1917.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
  Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; Pubmed=11259647;
RA Niernan W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,

```


RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Betty K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 4-phosphate (MEP) (by similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) ->
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE005865; AAK23892.1; -
 CC TIGR: CC1917; -
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom. 1.
 DR TIGRFAMS: TIGR00243; Dxr: 1.
 KW Isoprene biosynthesis; Oxidoreductase; NADP: Complete proteome.
 FT NP_BIND 13 20 NADPH (POTENTIAL).
 SQ SEQUENCE 399 AA; 41456 MW; B68D94C6D0A47C25 CRC64;

Query Match 45.3%; Score 899.5; DB 1; Length 399;
 Best Local Similarity 48.6%; Pred. No. 2.5e-56;
 Matches 191; Conservative 52; Mismatches 147; Indels 3; Gaps 2;

QY 2 RSLSTFGATGSGESTFEDLVMRKSGPEAFRTVALTNGRNIRRLAEMARLAKELAYTAHE 61
 DB 8 RRVVVGSTGSGTSLTSLFEESGAP--VQIATLTAGRNVERIEGARRRKPSLAVIDE 65
 QY 62 DCLPALREALAGTGTAVAGAAQAIAEADRPADMTSAITVGAAGLVGMRALKHGTAL 121
 DB 66 SRIDDLRAQIAGTGVAAAGADAVRDAAAGADWMSAIVGAGLAPVAAATGAVIAL 125
 QY 122 ANKESLVTAAGQLMRTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 181
 DB 126 ANKESLVCAAGPALATAKAGAGSVIPVDESHSAIFOVLOSECAHRSRLILTRASGGPFD 185
 QY 182 WSLERIRACTVAEAQAHPPMWSMGORISIDSASFNFKALELIERREFGEPPDIEAVHP 241
 DB 186 WDKAAMARATPEQAIHPPMWSMGAKISVDSATPMNKGLEIEISYLFATPEDVDVVIHP 245
 QY 242 OSIIVHAMVFCDDGLMAHLGAPADMRHAIGFALNMPGRGEVPAVIDLAQIASLTFOKPE 301
 DB 246 OSIIVHSLVYVGGSTLAQGGPDMRAPICAFAPMDRLPWPAPRDLAAYGGLTFESPVD 305
 QY 302 ERPPALRLARDVMAARGISGAAFNAKETAALDHFITAGRICGLDMAVVEETLAVG-STDP 360
 DB 306 ERPPALRLARDVMAARGISGAAFNAKETAALDHFITAGRICGLDMAVVEETLAVG-STDP 365
 QY 361 LFGKVPDALEEVYALMDHLARRAAEEAAGLRQK 393
 DB 366 LSVAESDAVETAMLIDGSARRIAAEVVAOKROR 398

RESULT 4
 ID DXR_ZYMO STANDARD: PRT: 388 AA.
 AC 09X5F2: 09RIA9:
 DT 30-MAY-2000 (rel. 39, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)

DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
 DE reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
 GN DXR.
 OS Zymomonas mobilis.
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
 OC Zymomonas.
 OX NCBI_Taxid=542;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-ATCC 31821 / ZM4 / CP4;
 RX MEDLINE=20461176; PubMed=11004410.
 RA Grollie S., Bringer-Meyer S., Salm H.;
 RT "Isolation of the dxr gene of *Zymomonas mobilis* and characterization
 RT of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase";
 RL FEMS Microbiol. Lett. 191:131-137(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 31821 / ZM4 / CP4;
 RA Lee H.J., Kang H.S.;
 RT Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 4-phosphate (MEP).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) ->
 CC -1- COFACTOR: REQUIRES A DIVALENT CATION; MAGNESIUM, MANGANESE OR
 CC COBALT.
 CC -1- ENZYME REGULATION: INHIBITED BY POSITRONICIN.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AJ250714; CAB60758.1; -
 CC EMBL: AF124757; AAD29659.1; -
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom. 1.
 DR TIGRFAMS: TIGR00243; Dxr: 1.
 KW Isoprene biosynthesis; Oxidoreductase; NADP; Magnesium; Manganese;
 KW Cobalt.
 FT NP_BIND 10 17 NADPH (POTENTIAL).
 FT CONFLICT 226 226 F -> Y (IN REF. 2).
 SQ SEQUENCE 388 AA; 41842 MW; 2E161B712089613F CRC64;

Query Match 44.1%; Score 875; DB 1; Length 388;
 Best Local Similarity 46.9%; Pred. No. 1.3e-54;
 Matches 183; Conservative 67; Mismatches 130; Indels 10; Gaps 3;

QY 2 RSLSTFGATGSGESTFEDLVMRKSGPEAFRTVALTNGRNIRRLAEMARLAKELAYTAHE 61
 DB 5 RYTVVVGATGSGTSLTSLDILIER--NDRYQVATLTANRNVKDLADAKRTNARRAYIADP 62
 QY 62 DCLPALREALAGTGTAVAGAAQAIAEADRPADMTSAITVGAAGLVGMRALKHGTAL 121
 DB 63 SLYNDLKELAGSSVDAAGADALVAAAMGADWMTAAIIGCGAKATLALIRKGTVAL 122
 QY 122 ANKESLVTAAGQLMRTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 181
 DB 123 ANKESLVASAGLIMIDAVREHGTLLPVDESHNAIFQCFPHNNDYVRRIITASGCPFT 182
 QY 182 WSLERIRACTVAEAQAHPPMWSMGORISIDSASFNFKALELIERREFGEPPDIEAVHP 241
 DB 183 TSLAEMATVYTPERAVOHPPMWSMGAKISIDSATPMNKGLEIEFHLFQIPLKFEFELVHP 242
 QY 242 OSIIVHAMVFCDDGLMAHLGAPADMRHAIGFALNMPGRGEVPAVIDLAQIASLTFOKPE 301

```

Db 243 QSVHSMWEYLDGSIILQIGSPDMKRTPIGHTLAMPKRETTAESIDETKLMQMFEDDY 302
Oy 302 ERPALRLARDVMAARGLSGAFFNAKEIALDHFIAIGIGFLDMAAVEETLACSTPL 361
Db 303 ERFPALRLAMSTISGSGARRPVMANAEIYAVALDKKIGFLDIKIVKTL-----DHY 357
Oy 362 FGKVPDALEEVYLANDHLARRAAEAGLRQ 391
Db 358 TPATPSSLEDFVAIDNEAR---IOAAALME 384

RESULT 5
DXR_RALSO STANDARD; PRT; 393 AA.
ID DXR_RALSO STANDARD; PRT; 393 AA.
AC 08X215;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xyulose-5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR RSC1410 OR RS05282.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; proteobacteria; beta subdivision; Ralstonia group.
OC Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM1100;
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangelot S.,
RA Ariat M., Billault A., Brotter P., Camus J.C., Cattolico L.,
RA Chandler C., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RT Nature 415:497-502(2002).
RL
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xyulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + (+)
-1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AL646064; CAD15112.1;
DR InterPro: IPR003821; DXP_reductoisomase.
DR Pfam: PF02670; DXP_reductoisom; 1.
DR TIGRPFAMS: TIGR00243; Dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NF_BIND 7 NADPH (POTENTIAL).
SQ SEQUENCE 393 AA; 41763 MW; 5A995A7C8F32E5D CRC64;

Query Match 40.2%; Score 797.5; DB 1; Length 393;
Best Local Similarity 43.6%; Pred. No. 3.6e-49;
Matches 170; Conservative 73; Mismatches 134; Indels 13; Gaps 5;

Oy 1 MRSISTGAGSGTGFEDLVMRKSGPEARTVALTGGRNIRRLAENARLAKALAVTAH 60
Db 1 MMRITVGAAGSGTSDTLDVVRH--PDYRVFALTANTQVDKLAALCRVFRPMAAVGS 58
Oy 61 EDCIAPALRELA--GCTGEVAGAGQAIAEAADR-ADMTSATVAGAGVPGMALKHGR 117

```

```

Db 59 ATAAEVLRDQGAERATGIDIRFGEPALEEAAPHDCOAVMAIYGAAGLRPTLAVRAGK 118
Oy 118 TLALANKESVYTAGQLMRTAQENGATILPVDSHSAVFQALAGEDPAC---VERVITTA 174
Db 119 RVLANKREKALVMSGLAFMDAARGATVLPIDSEHNALFQCLPQORPSFGVGARIYVTA 178
Oy 175 SGGEPRDMSLERIRACVYAEQAHPNMSGQRISIDSAFMENKALELLETRFEFGFEPDR 234
Db 179 SGGEPRTRAVETLAEVFPDQACAPHNWVGRIKISVDSATMANKGLEVEAHMLFNVEVER 238
Oy 235 IEAVVHPQSYVHAAMVFCDDGLMHLGPADMRHAIIGFLNMPRGGEVPAVIAQLASTL 294
Db 239 LEVLIHPQSVIHSWVAIDGSLVLAQLGNPDMRTPIAGLAPERIEAGVPLDLAATGLT 298
Oy 295 TFOKDEERFPALRLARDVMAARGLSGAFFNAKEIALDHFIAIGIGFLDMAAVEETLA 354
Db 299 AFEAPDLHRFPCLALFALRAGTAPRVLANAEVNEARLQGRIRRTETIAAIVGDTLA 358
Oy 355 GVSTDPLEFGKVPDALEEVYLANDHLARRAAE 384
Db 359 RTAIGP-----ADSIDTVEFAADQAARRRAE 383

RESULT 6
DXR_PSEAE STANDARD; PRT; 396 AA.
ID DXR_PSEAE STANDARD; PRT; 396 AA.
AC 09KGD6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR PA3650.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE-20487113; PubMed-11034300;
RA Allincoek B., Hintz M., Sanderbrand S., Wiesner J., Beck E.,
RA Jomaia H.;
RT "ools for discovery of inhibitors of the 1-deoxy-D-xyulose
5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with
enzymes from the pathogenic bacterium Pseudomonas aeruginosa.";
RT FEMS Microbiol. Lett. 190:329-333(2000).
RL
RN
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huinagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xyulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
-1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- ENZYME REGULATION: INHIBITED BY POSITROMYCIN AND 3-(N-ACETYL-N-
HYDROXYAMINO)-PROPYLPHOSPHONIC ACID (PR-900098).
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

[illegible]

OX NCBI_TaxID-65699;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN-22491 / Serogroup A / Serotype 4A;
 RA MEDLINE-20222556; PubMed-10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skellon J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RA "complete DNA sequence of a serogroup A strain of *Neisseria*
 RT meningitidis 2291.";
 RL Medline 404:502-506(2000).
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 CC 4-phosphate (MEP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AL162752; CAB83399.1;
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr: 1.
 KM Isoprene biosynthesis; Oxidoreductase; NADP: Complete proteome.
 FT NP_BIND 9 NADPH (POTENTIAL).
 SO SEQUENCE 394 AA; 41912 MW; B5D361B224806D58 CRC64;
 Query Match 37.5%; Score 743.5; DB 1; Length 394;
 Best Local Similarity 43.0%; Pred. No. 2.3e-45;
 Matches 169; Conservative 58; Mismatches 151; Indels 15; Gaps 5;
 OY 4 LSTIGANGSGESTFDLMKRGGEAFRTVALTGCGRNIRKAEAKALKALATTAHEDC 63
 DB 6 LTIIGSTGSGESTLDDVSRH--PEKFRFVALAGHKOVEKLAACQFTFHEPRTAVVADAEH 63
 OY 64 LPALREMLA--GTGTEVAGGAQAIAEAADR-PADMTSAIAGAGLVPGMALKHGRTLA 120
 DB 64 AARLEALLKRGATVQLHGAQALVDVASADEVSGVCAITGAVGLPSALAAOKGRTIY 123
 OY 121 LANKESLVTAQGLLMKRTAENGATILPVDESHSAVFOALAGEDTA-----CVERVIITAS 175
 DB 124 LANKETLVVSGALFEMETARANGAVALPVDSEHNAVFOVLPDRTYGRNENHGIASITLITAS 183
 OY 176 GGPRDMSLEIRACTVAEQAHPNMSGORISIDSAMFKALELLETREFFGEPDR 235
 DB 184 GGPELADLNTFDSITFDQAVKHPNMRGRISVDSATMMKRGLELLEAHLFNCPPDKL 243
 OY 236 EAVVHPOSIYAVAGFGCDGLMAHLGPAIDMRHAGFALNMPGGEVPAARDLQIASLT 295
 DB 244 EYVTHPOSVIHSWRYKRDGSLADLGNPDMPRTATYCLGPERIDSGVGLDIFALSALT 303
 OY 296 FQKDEERFPALRLARVMAAGLSGAFAVNAKEIALDHFAGRIIGFLDMAVAVEETLAG 355
 DB 304 FQKDFDFRFPCLAYEMANAGAAPCVLNANANAAVAFLDGOIKFTDIAKTVAHCLSQ 363
 OY 356 VSDPLEFGKVPDALEEVLANDHLLARRAAEEELAG 388
 DB 364 DFDSDGI-----GDIGGLADADARTRAQARAFTIG 391
 RESULT 9

DXR_ARATH
 ID DXR_ARATH STANDARD; PRT; 477 AA.
 AC O9XFS9: O9M6U2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplast precursor
 DE (RC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate
 DE reductoisomerase).
 GN DXR OR AT5G62790 OR MGB2.11.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RA Campos N., Lois L.M., Cunillera N., Carretero L., Alameda I.,
 RA Hoefli J.-F., Pale-Grosdemange C., Rohmer M., Ferrer A., Boronat A.;
 RT "Isolation and characterization of a cDNA from Arabidopsis thaliana
 RT encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first
 RT committed enzyme of the non-mevalonate pathway for isoprenoid
 RT biosynthesis.";
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN [3]
 RP SEQUENCE OF 72-477 FROM N.A.
 RC STRAIN-cv. Columbia;
 RA Schneider J., Mueller C., Zeidler J., Lichtenhaler H.K.;
 RT "Cloning and heterologous expression of a cDNA coding for 1-deoxy-D-
 RT xylulose-5-phosphate reductoisomerase of Arabidopsis thaliana.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 CC 4-phosphate (MEP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF148852; AAF73140.1;
 DR EMBL: AB009053; BAB10848.1;
 DR EMBL: AJ242588; CAB43344.1;
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr: 1.
 KM Isoprene biosynthesis; Oxidoreductase; NADP; Chloroplast;
 KW Transic peptide.
 FT TRANSIT 1
 FT CHAIN 1
 FT NP_BIND 86 93
 FT SEQUENCE 477 AA; 51963 MW; C06A455AB73ACA7C CRC64;
 SO

CC RA da Silva A.C.R., da Silva F.R., Silva W.J., de Souza A.A.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchino M.H.,
Valada H., Van Sluyms M.A.I., Verjovski-Almeida S., Vettore A.L.,
Zago M.A., Zatz M., Meidanis Y., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -I- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of l-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP) (By similarity).
CC -I- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADPH(+)
= l-deoxy-D-xylulose 5-phosphate + NADH.
CC -I- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.

CC CC

-I- SIMILARITY: BELONGS TO THE DXR FAMILY.

CC CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb-sib.ch/announce/or send an email to license@lsb.sib.ch). -----

DR EMBL; AE003942; AACF8358.1; -
DR InterPro; IPR003821; DXP_redoxsomase.
DR Pfam; PF02670; DXP_reductoisom; 1.
DR TRIGRAMS; TIGR00243; dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 11
SQ SEQUENCE 396 AA; 42076 MW; 1AB5FE95748FD97 CRC64;

Query Match 37.3%; Score 740; DB 1; Length 396;
Best Local Similarity 42.4%; Pred.No. 4e-45;
Matches 165; Conservative 60; Mismatches 154; Indels 10; Gaps 5;

OY 1 MRSLISFGATSGISESTVDLYMKRGKGPAPFRVALTGGRNIRRLAEAKALAYTAH 60
::: : | ||||| :: : | : : : : :
5 IRNAVLCGATGSIGAALDVLARR--PPGFVHSLLAAGQRDALCALCHTPRDHAVIGD 62
::: : | ||||| : | : | : : | ||||| : :

OY 61 EDCLPALREAL--AGGTGEVGAQAQLA--AADRPADMYSIVYGACLVPGMRALKHGK 117
::: : | ||||| : | : | : : | ||||| : :

D b 63 ATVTTLTDGLMAGLAIKATKAAGEALLBELVASTCDTVAAIYAAGLSHTILAARKG 122
::: : | ||||| : | : | : : | ||||| : :

OY 118 TLALANKESLTAGOLLMTAOENGATILLPVDSHASVFQALAGEDTCACVERVIITASGG 177
| ||||| | ||||| : | : | : | ||||| : | : |||||
D b 123 RLLLANKESSLVLAGMLLRKASISCAETIIPIIDSEHNIAIFOCLESRSTTGVRITPLAAG 182
| ||||| | ||||| : | : | : | ||||| : | : |||||

OY 178 PEROMSLERINACYAAEQAHPNMSMGORISIDSAMNKALELETPEFFEFEDRIEA 237
| ||||| : | : | ||||| ||||| : | ||||| : : |:
183 PFRGNRNTRLAKKTPTQQMAHPWTSMGKRISDVSTYLNNKGILEVEANHLGFLPSQSDIV 242
| ||||| : | : | ||||| ||||| : | ||||| : : |:

OY 238 VVHQSIYHAHVGCDDGIIMNLHGPSADRHHAIIGEAILMPNGREVENVARIDLIALQSISTLF 297
: | ||||| : | : | ||||| : | : | : | ||||| : : |:
OY 243 LVHQSVLHSVLEFI DGSTLSQLSPDMRKRTLAVLGSNPERIGSSVPGLDKMKHRDFE 302
| ||||| : | : | ||||| : | : | : | ||||| : : |:

OY 298 KPDEBRPALKLARDVMARGSCAFNAAKEIALDHFIAGRIGFLDMAAVEETLAGVS 357
: | ||||| : | : | ||||| : | : | : | ||||| : : |:
D b 303 KPDTEFSCLRALARDAWGCTGAPAVLIANAANEIASFAFLQGRIIGFLTIPALLEHAL---T 359
| ||||| : | : | ||||| : | : | : | ||||| : : |:

OY 358 TDPLEGVKDPALEEYLDMDHLARRAAEEA 386
| ||||| : | : | ||||| : | : | : | ||||| : : |:

D b 360 TLPRK--EADTLETLILTYDTETRITTAA 386
| ||||| : | : | ||||| : | : | : | ||||| : : |:

RESULT 11
D XR_ECOS7 STANDARD; PRt; 398 AA.
AC Q8XBY1:

TT 15-JUN-2002 (Rel. 41, Created)
TT 15-JUN-2002 (Rel. 41, Last sequence update)
TT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 1-deoxy-D-xylulose 5-phosphat reductoisomeraze [EC 1.1.1.267] (DXP

-Qy	228	FGEPEDLTAEVHHOSIVHAMGEFCGGMAHLGAPDMHATGFLNMPGREGVEVPVARID	287
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	239	FNASASOMEVLLIHPOSTHSHVRIQDSVLAQGEPMKTRPLAHWAMPNRYNSGVKFLD	298
Qy	288	LAQIATSLTFQKPEEERFPAIRLARADYWARGLSGAFAFNARKEIALDHFIAGRIGFLDMAA	347
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	299	FCRLSATLTPAAPDYDRPCILKAMEAFEGQQAATTATLNMNANETLVAAFLAQIRFTDIAA	358
Qy	348	VVEETLAGVSTDPLEFGKVPDALEEVLAAMDLARRAAEE	385
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	359	LNLSVL-----EKKMDMREPOCVDDVLVSVDASAREVARK	391
RESULT 12			
DXR	ECOLI	STANDARD:	PRT: 398 AA.
ID	DXR_ECOLI		
AC	P45568; P77209;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	1-deoxy-D-xylose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).		
GN	DXR OR B0173.		
OS	Escherichia coli.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia		
OX	NCBI_TaxID=562;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.		
RC	STRAIN-K12 / W3110;		
RC	MEDLINE=98374274; PubMed=9707569;		
RA	Takahashi S., Kuzuyama T., Matenabe H., Seto H.;		
RT	"A 1-deoxy-D-xylose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis."		
RL	Proc. Natl. Acad. Sci. U.S.A. 95:9879-9884(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12 / MG1655;		
RC	MEDLINE=97426517; PubMed=9278503;		
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,		
RA	Riley M.K., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,		
RA	Gregor J., Davis N.W., Kirkpatrick J.A., Goeden M.A., Rose D.J.,		
RA	Mau B., Shao Y.;		
RT	"The complete gene sequence of Escherichia coli K-12."		
RL	Science 277:1453-1474(1997).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12 / W3110;		
RA	Takemoto K., Mori H., Mureyama N., Kataoka K., Yano M., Itoh T.,		
RA	Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,		
RA	Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;		
RT	"Systematic sequencing of the Escherichia coli genome: analysis of the 4.0 - 6.0 mln (189,987 - 281,416bp) region."		
RL	Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,		
RA	Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurd O.,		
RA	Leshkari D., Lew H., Lin D., Namath A., Oelner P., Roberts D.,		
RL	Davis R.W.;		
RP	Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.		
RN	[5]		
RP	SEQUENCE OF 1-49 FROM N.A.		
RC	MEDLINE=93077430; PubMed=1447125;		
RA	Yamanaka K., Ogura T., Miki H., Hiraga S.;		
RT	"Identification and characterization of the smda gene, a suppressor of the mukB null mutant of Escherichia coli."		
RL	J. Bacteriol. 174:7517-7526(1992).		
RN	[6]		
RP	IDENTIFICATION.		
RC	MEDLINE=96032851; PubMed=7567469;		

```

QY      118 TLTANESLVTAGQLMRTAEOENGATLPDSSESAVFAOL-----AGEDTACY 167
Dy      119 TLTANESLVTGRLFDMAVKQSAOALLPYDSEHNAITFQSPDPDTIGNLGADDEONGV 178
QY      168 ERVITTAAGSGFRBMSLERACTVAEAQAHPNMSGORISIDSASMPNKALELETETEF 227
Dy      179 VSILLTGSGGFRETPLDLATMTMPDQCRRPNNMSGKRISSVDSTAMNKGLEYEARML 238
QY      228 FGFEPRLEAVNHQSIYHANVGCDDGLAAHLGPADKRRAHIGFLANMPGGEVYVARID 287
Dy      239 FNASAQMEVIHIPOSVSHSMVRQDGSVLAOLGPDMDRTPIAHTMAPNRNSGVKPLD 298
QY      288 LAOIASITFOQPDEREPALRLAROVMAARGLSGAFAAFKAIEALDHTIAGRIGELDMA 347
Dy      299 FCKISALTFAAPDVDRYCLKLAMEAFPGOGAATTALMANEITYAATLAQOIREFTDIAA 358
QY      348 VVEETLAGVSTDPLEGKVDALEEVLDAMHLARRAAEE 385
Dy      359 LNLSQL-----EKMDREPOCVDVLSDVANAREVARK 391

RESULT 13
DxR_YERPE STANDARD: PRT; 398 AA.
AC Q8ZH62;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase).
GN DXP OR YP01048.
OS Versinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Versinia
OX NCBI_Taxid=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CO-92 / Biovar Orientalis;
RX MEDLINE-21470413; PubMed-11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,
Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
Baier S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
Chillingworth T., Cronin A., Davies R.K., Davis P., Dougan G.,
Felwell T., Hamlin N., Holroyd S.J., Jagels K., Karlyshev A.V.,
Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
= 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
Dy EMBL; AJ414146; CAC98980.1; -
DR InterPro; IPR003821; DXP_reductoisomase.
DR Pfam; PF02670; DXP_reductoisom; 1.
DR TIGRFAMs; TIGR00243; dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 7 14 NADH (POTENTIAL).
SQ SEQUENCE 398 AA; 43115 MW; 69B1PC9ED016D057 CRC64;
```


Query Match 36.5%; Score 725; DB 1; Length 398;
 Best Local Similarity 40.7%; Pred. No. 4,6e-44;
 Matches 162; Conservative 69; Mismatches 147; Indels 20; Gaps 5;

QY 1 MRSISFGATSGISGSTDVLMRKGGPEAFRTVALTGGNRIRLAEMARLKAELAVTAH 60
 D 1 MKOLITIGSTGISTNSTLSV--RANPELFKVTALVAGRNREMAOCCLEFSPPRYAASD 58
 QY 61 EDCLPALREALA--GTGTEVAGGAOAIAE--AADRPADMTSAIGAALVPGMRALKGR 117
 D 59 EHSKSLRLLLAEQSDTEVSGETACELALDDVDVMAIVGIAELPSTLAIRAGK 118
 QY 118 TLALANKESLVTAQOLMRTAENGATILPVDSHSAVFQAL-----AGEDTACV 167
 D 119 QVLLANKESLITCGLFNDEVKRSRAQLLPIDSEHNAIFQSLPERIQRLQSYSLNMGV 178
 QY 168 ERVITTAGSGGFPDMSLERIRACTVAEAOAHPNMSMGORISIDSAMFNKALELIEFREF 227
 D 179 SKILLTSGGPFRTPTLSQFSDVTPDQCAHPNMSMGRIKISVSTVMNKGLEYIEARWL 238
 QY 228 FGPEPRIEAVVHPOSIVAHVAGCGDGLMAHLGPADMRHAIIGFALNPGGEVVARID 287
 D 239 FNASAEQLEVLHPQSVHSHVRIHDSILQMGTPDKRTIARAHMAYPMKVSQVAPLD 298
 QY 288 LAQIASITFORPDERPALRLARDVMAARGLSGAAPNAKEIALDHPFIAGRIGFLDMAA 347
 D 299 FCKVGLFTTPDYQRYRCLKLAIDACNAGOAATFALNAANEISVMAFLDSKIRTDIEV 358
 QY 348 VVEETLACVSTDPFLGKVPDALEVLAMDLARRAAE 385
 D 359 INRTVEG-----DLSEPTVEEVLVIDRARDVAAQ 391

RESULT 14
 DXR_HAEIN STANDARD; PRT; 397 AA.

AC P44055;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-deoxy-D-xyululose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxyululose-5-phosphate reductoisomerase).
 GN DXR OR H10807.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M., Grehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
 RT Science 269:496-512(1995).
 RL
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction of 1-deoxy-D-xyululose-5-phosphate (DXP) to 2-C-methyl-D-erythritol 4-phosphate (MEP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) -> 1-deoxy-D-xyululose 5-phosphate + NADPH.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second step.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

DR EMBL: U32763; AAC22466.1;
 DR TIGR: H10807;
 DR InterPro: IPR003821; DXP_reductoisomase.
 DR Pfam: PF02670; DXP_reductoisom: 1.
 DR TIGRFAMs: TIGR00243; Dxr: 1.
 KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 9
 FT NADPH (POTENTIAL).
 SQ SEQUENCE 397 AA; 43667 MW; A6E9B24CC6A1CEAA CRC64;

Query Match 36.1%; Score 717; DB 1; Length 397;
 Best Local Similarity 40.2%; Pred. No. 1.7e-43;
 Matches 159; Conservative 72; Mismatches 145; Indels 20; Gaps 5;

QY 2 RLSISFGATSGISGSTDVLMRKGGPEAFRTVALTGGNRIRLAEMARLKAELAVTAH 61
 D 4 QNIVILGSGISGISTLSV--ENNPOKYHAFALVGGKNEAMEEQCIKRRPRAALDDV 61
 QY 62 DCLPALREALG--TGTEVAGGAOAIAEADRP--ADWTMSAIVGAGLVPGMRALKHGR 118
 D 62 NAAKILREKLLAHNIHPVELVLAGRAICELAHNPADQIMASIVGACGLPLTSVAVKGR 121
 QY 119 LALANKESLVTAQOLMRTAENGATILPVDSHSAVFQALG-----DTACVE 168
 D 122 VLLANKESLVTCGOLFIDAVANYSKLLLPVSEHNAIFQSLPPPAOEKIGCPSELGVS 181
 QY 169 RVITTAGSGPFDMSLERIRACTVAEAOAHPNMSMGORISIDSAMFNKALELIEFREF 228
 D 182 KIILLTSGGPFRTPTLSQFSDVTPDQCAHPNMSMGRIKISVSTVMNKGLEYIEARWL 241
 QY 229 FGPEPRIEAVVHPOSIVAHVAGCGDGLMAHLGPADMRHAIIGFALNMPGGEVVARIDL 288
 D 242 NASAEQLEVLHPQSVHSHVRIHDSILQMGTPDKRTIARAHMAYPMKVSQVAPLD 301
 QY 289 AQIASITFORPDERPALRLARDVMAARGLSGAAPNAKEIALDHPFIAGRIGFLDMAA 348
 D 302 FKIELFTFEPDNRPFLKLAIDAFAGGYATTANANANEIVAOFLDRDIGFMDIAKI 361
 QY 349 VVEETLACVSTDPFLGKVPDALEVLAMDLARRAAE 384
 D 362 NKTIERISPYTI-----QNTDVLVEIDAQAREIAK 392

RESULT 15
 DXR_PASMU STANDARD; PRT; 405 AA.

AC P57985;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-deoxy-D-xyululose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxyululose-5-phosphate reductoisomerase).
 GN DXR OR PM1988.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_Taxid=747;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.-J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction of 1-deoxy-D-xyululose-5-phosphate (DXP) to 2-C-methyl-D-erythritol

CC 4-phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006235; AAK04072.1; -
DR InterPro: IPR003821; DXR_red isomerase.
DR Pfam: PF02670; DXR_red isom; 1.
DR TIGRfams: TIGR00243; Dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 13 20 NADPH (POTENTIAL).
SQ SEQUENCE 405 AA; 44207 MW; 1CFF7B44847F8BF7 CRC64;

Query Match 36.1%; Score 716; DB 1; Length 405;
Best Local Similarity 39.3%; Pred. No. 2e-43;
Matches 156; Conservative 74; Mismatches 147; Indels 20; Gaps 5;

QY 1 MRSLSIFGATSGISGESTPDLVNRKSGPEAFRTVALTGGRNIRRLAMARAKLAELAVTAH 60
DB 7 MKRIYILGSGTSGISGTSLSVTH--NPKYQVAFVAGRNVELMFQOCLTFQPSFALDD 64

QY 61 EDCLPALREALAG--TGTEVAGGAQAIAEADRP--ADWTMSAIVGAAGLVPGMRALKHGR 117
DB 65 DVAAMKLAELKAKHQSQTIVLAGQAICELAHPEADMMAIVGAAGLPLTSAVKAKR 124

QY 118 TLALANKESLVTAQQLMRTAQENGATILPVDESHSAVPQALAGE-----DTACV 167
DB 125 RVLANKKALVTCGQLFIDAVRESQQLPVDSEHNAIFQSLPPEAQRQIGCPLELSEGI 184

QY 168 ERVITTAGSGPFRDMLEIRACTVAEAAHNMWSMGORISIDSASMEKALELITREF 227
DB 185 SKIVLGSQGPFRYPPELFQGITPAQAAVAHPWMSGKKISVDSATVMNKGLEYIEARWL 244

QY 228 FGFEEDRIEAVVPOSIVAMVFCDDGLMAHIGPADMRHAIGFALNMPGRGCEVPVARI 287
DB 245 FNASAEEMEVIIHPOSIIHSMRYIDGSVIAQMGNDMRTPIAETMAYPSRTVAGVEPLD 304

QY 288 LAQIASLTFQKPDDEERFALRLARDVMAARGLSGAAFNAKETALDHTAGRTIGFLDMA 347
DB 305 FYQLNGLTFTIEPDYGRPCIKLAIDAFSAGQYATTMANANEIAVASFLDNKIKFTDIAR 364

QY 348 VVEETLAGVSTDLFGKVPDALAEVYLAAMDHLARRAE 384
DB 365 LMQLVYSKILQPKI-----HCIEDVLEVDKKARELSQ 396

Search completed: April 15, 2003, 13:57:00
Job time : 26 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:29 ; Search time 88 Seconds

(without alignments)
922.529 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984
Sequence: 1 MRSLSIFGATSGISGFEDL.....MDHLARRAEAGHROOKR 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SEPREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	994	50.1	397	16	08UC86
2	754	38.0	472	10	09SP64
3	747	37.7	473	10	09FTNO
4	746	37.6	472	10	09FX27
5	745	37.6	475	10	0947C3
6	740	37.3	473	10	08W250
7	715	36.0	474	10	09M4W4
8	684.5	34.5	382	16	08RA28
9	665.5	33.5	386	2	09A4D7
10	568	28.6	390	16	08R622
11	519.5	26.2	488	5	09G693
12	181.5	9.1	115	2	09EX11
13	136.5	6.9	2555	11	091Z66
14	136.5	6.9	2561	11	08VIF5
15	133.5	6.7	412	16	09A4V4
16	130.5	6.6	981	16	08XS09

17	128.5	6.5	442	16	098FT2	098ft2 rhizobium 1
18	124	6.2	439	16	08XZ88	08xzt8 ralstonia s
19	123.5	6.2	1731	16	08U8M4	08u8m4 agrobacteri
20	122.5	6.2	3247	12	065553	065553 bovine herp
21	121.5	6.1	585	16	09ADP2	09adp2 streptomyc
22	121	6.1	1041	2	093C90	093c90 actinomadr
23	119.5	6.0	5435	2	0914X2	0914x2 streptomyc
24	119	6.0	538	16	050442	050442 mycobacteri
25	118.5	6.0	4881	2	09S0R3	09s0r3 streptomyc
26	118	5.9	437	2	09RAM9	09ram9 methylobact
27	117	5.9	520	16	09A436	09a436 caulobacter
28	115.5	5.8	434	16	08YD51	08yds1 brucella me
29	115.5	5.8	9477	2	0914X3	0914x3 streptomyc
30	115	5.8	1728	2	054591	054591 amycolatops
31	115	5.8	3190	2	093NW7	093nw7 streptomyc
32	114.5	5.8	3970	2	093H18	093h18 streptomyc
33	113.5	5.7	7257	2	0918C7	0918c7 polyangium
34	113.5	5.7	7257	2	09K1Z7	09k1z7 polyangium
35	112	5.6	711	16	P96855	P96855 mycobacteri
36	112	5.6	951	2	068078	068078 rhodobacter
37	112	5.6	1472	16	091214	091214 streptomyc
38	111.5	5.6	985	2	092N16	092n16 pseudomonas
39	111.5	5.6	1907	2	093H19	093h19 streptomyc
40	110.5	5.6	429	16	09RD83	09rd83 streptomyc
41	110.5	5.6	554	16	092T76	092t76 rhizobium m
42	110.5	5.6	941	16	092IWS	092iws streptomyc
43	110.5	5.6	1099	16	09A4D6	09a4d6 caulobacter
44	110.5	5.6	1281	16	091211	091211 pseudomonas
45	110.5	5.6	1293	2	005170	005170 streptomyc

ALIGNMENTS.

RESULT 1

ID	Q8UC86	PRELIMINARY:	PRT:	397 AA.
AC	08UC86;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	1-deoxy-D-xylose 3-phosphate reductoisomerase.			
GN	DXR OR ATU2617 OR AGR_C_4736.			
OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Rhizobiaceae; Rhizobium.			
OX	NCBI_TaxID=176299;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21608550; PubMed=11743193;			
RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,			
RA	Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,			
RA	Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,			
RA	Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,			
RA	Kutyavin T., Levy R., Li M.-J., McCelland E., Palmeri A.,			
RA	Raymond C., Rouse G., Senechinnachak C., Wu Z., Romero P., Gordon D.,			
RA	Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,			
RA	Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,			
RA	Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,			
RA	Nester E.W.;			
RT	"The genome of the natural genetic engineer Agrobacterium tumefaciens			
RT	C58. ";			
RT	Science 294:2317-2323(2001).			
RT	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21608551; PubMed=11743194;			
RA	Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,			
RA	Qurollio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,			
RA	Hounell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,			
RA	Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,			
RA	Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,			
RA	Cielo C., Slater S.;			
RT	"Genome sequence of the plant pathogen and biotechnology agent			

RT Agrobacterium tumefaciens C58.
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009208; AAL43593.1;
 DR EMBL: AE008174; AAK88334.1;
 KW Isomerase; Complete proteome.
 SO SEQUENCE 397 AA; 42303 MW; 0A24C408C2D47094 CRC64;

Query Match 50.1%; Score 994; DB 16; Length 397;
 Best Local Similarity 50.4%; Pred. No. 1.3e-54;
 Matches 197; Conservative 64; Mismatches 126; Indels 4; Gaps 1;

QY 2 RSLIFGATGSGISTPDLVNRKGGPEAFRTVALTGGRNIRRLAEMARLAEAVTAHE 61
 Db KRLTILGSGISGTITLVDVROIGRGDFEETMALTGAGNIALLAQAEFGAQLAVTAD 68
 QY 62 DCLPALREALGTEVAGCAQAIAEADRPADWTMSAIVAGAGLVPMRALKRGRTAL 121
 Db 69 DKYEALKALAGTGIKVAAGAAAGLEAASMDAGVMAIACITPGIAPLTLLARAGDIAL 128
 QY 122 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 181
 Db 129 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 188
 QY 182 WSLERIRACTVAEAOAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIRAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 248
 QY 242 QSIYHAWVFCDDGGMALHGPADMRHAIGFALNMPGRGEVVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDSYIAQIGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFAPRLARVMAARGLSGAFAFNKAEIALDHFIAIGTGLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFHAFVAGGIGGLDMAEIVTDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALAEIATK 395

RESULT 2

Q9SP64 PRELIMINARY; PRT; 472 AA.

AC Q9SP64; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 1-deoxy-D-xylose-5-phosphate reductoisomerase.
 GN DKR1.
 OS Artemisia annua (Sweet wormwood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; easterids II; Asterales; Asteraceae; Asteroidae;
 OC Anthemideae; Artemisia.
 OC NCBI_TaxID=35608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. YUGOSLAVIA; TISSUE=ROOT;
 RA Wobbe K.K., Souret F.F., Shore K.A., Weathers P.J.;
 RT "Artemisia annua 1-deoxy-D-xylose-5-phosphate reductoisomerase (dxr)
 RT mRNA";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF182287; AAD56391.2;
 DR InterPro: IPR003821; DXP_reductisomase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFams: TIGR00243; Dxr; 1.
 KW Isomerase.
 SO SEQUENCE 472 AA; 50740 MW; D52023C09D475675 CRC64;

Query Match 38.0%; Score 754; DB 10; Length 472;
 Best Local Similarity 41.1%; Pred. No. 1.7e-39;
 Matches 164; Conservative 77; Mismatches 136; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGISTPDLVNRKGGPEAFRTVALTGGRNIRRLAEMARLAEAVTAHE 61
 Db KRLTILGSGISGTITLVDVROIGRGDFEETMALTGAGNIALLAQAEFGAQLAVTAD 68
 QY 62 DCLPALREALGTEVAGCAQAIAEADRPADWTMSAIVAGAGLVPMRALKRGRTAL 121
 Db 69 DKYEALKALAGTGIKVAAGAAAGLEAASMDAGVMAIACITPGIAPLTLLARAGDIAL 128
 QY 122 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 181
 Db 129 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 188
 QY 182 WSLERIRACTVAEAOAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIRAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 248
 QY 242 QSIYHAWVFCDDGGMALHGPADMRHAIGFALNMPGRGEVVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDSYIAQIGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFAPRLARVMAARGLSGAFAFNKAEIALDHFIAIGTGLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFHAFVAGGIGGLDMAEIVTDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALAEIATK 395

RESULT 3

Q9FTNO PRELIMINARY; PRT; 473 AA.

AC Q9FTNO; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 1-deoxy-d-xylose-5-phosphate reductoisomerase.
 GN P0005A05.19 OR P0482C06.2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0005A05";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0482C06";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002863; BAB16915.1;
 DR EMBL: AP002845; BAB78606.1;
 DR InterPro: IPR003821; DXP_reductisomase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFams: TIGR00243; Dxr; 1.
 KW Isomerase.
 SO SEQUENCE 473 AA; 51473 MW; 9933D9D8C3D1FC49 CRC64;

Query Match 37.7%; Score 747; DB 10; Length 473;
 Best Local Similarity 41.4%; Pred. No. 4.7e-39;
 Matches 165; Conservative 74; Mismatches 138; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGISTPDLVNRKGGPEAFRTVALTGGRNIRRLAEMARLAEAVTAHE 61
 Db KRLTILGSGISGTITLVDVROIGRGDFEETMALTGAGNIALLAQAEFGAQLAVTAD 68
 QY 62 DCLPALREALGTEVAGCAQAIAEADRPADWTMSAIVAGAGLVPMRALKRGRTAL 121
 Db 69 DKYEALKALAGTGIKVAAGAAAGLEAASMDAGVMAIACITPGIAPLTLLARAGDIAL 128
 QY 122 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 181
 Db 129 ANKESLVTAAGLMTKTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGGPPRD 188
 QY 182 WSLERIRACTVAEAOAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIRAHNPMNSMGORISIDSAMFNKALELLETREFEFGEDRIEAVVHP 248
 QY 242 QSIYHAWVFCDDGGMALHGPADMRHAIGFALNMPGRGEVVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDSYIAQIGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFAPRLARVMAARGLSGAFAFNKAEIALDHFIAIGTGLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFHAFVAGGIGGLDMAEIVTDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALAEIATK 395

[illegible]


```

Db 240 VHPDSVHSHWEYVDGSIIMGVANKMPTILVAFYPEREYNSINFLDIKNNLT 299
QY 297 QKPDREPARLRADVAAGLSGAFNAKETAALDHFINGRIGELVMAVETLAG- 355
Db 300 EADBRKVLKJDLAVRACRGDTMPTVFNASNAIVELFMKKQIKELDIYRIEEMASH 359
QY 356 ---VSTDPLFGKVPDALEVLAMDLARRAAE 385
Db 360 QVLTNTD-----NALNVIKEVDREIRKKVRE 386

RESULT 11
096693 PRELIMINARY; PRT; 488 AA.
AC 096693;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 1-deoxy-D-xytlose 5-phosphate reductoisomerase.
GN DXR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RA Altincsek B., Sanderbrand S., Wiesner J., Jomaa H.;
RT "dxr as a potential target for antimalarial drugs.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF111813; AAD03739.1;
DR InterPro: IPR003821; DXP_reductoisomase.
DR TrEMBL: PF02670; DXP_reductoisom; 1.
DR TrEMBL: TIGR00243; Dxr; 1.
KW isomerase.
SQ SEQUENCE 488 AA; 55756 MW; 4E280C81CDFAD3EF CRC64;

Query Match 26.2%; Score 519.5; DB 5; Length 488;
Best Local Similarity 30.5%; Pred. No. 8.5e-25;
Matches 124; Conservative 88; Mismatches 167; Indels 27; Gaps 8;

QY 3 SLSIFGATSGESPFEDLVMRKGPE-AFTVALTGSNIRRLAEMALAEAVTAE 61
Db 79 NVALFGSGISGTNALNIRCKNIENFVKALVNSVELLEQAEFLPEYLCHDK 138
QY 62 DCLPALREALGTGE--VAGQAQIAE-AADRPADWTSAYGAAGLVGMRALKHGR 117
Db 139 SYEELKELVKNIKDYRIILCGEGMKECSSNSIDKIVIGIDSFQGLYSFMVAIMNK 198
QY 118 TLALANKESLYTAGOLMRTAQ-ENGATILPVDSHSAVFPALAGE---DTAC----- 166
Db 199 IVALANKESIVSAGFEFLKLNIRKNAKIIPVDSHSAIFQCLDNKVKLTKCLQDNFSK 258
QY 167 ---VERVITASGPFPMSSLERIACVAAEQAHPMNSMORISIDSASFENKALELIE 223
Db 259 INNINKIFLCSGGPFQULMDELKNYSENALKHPKKMKKTKITIDATYMNKGLYEIE 318
QY 224 TRFEGEPRIEAVVHQSIIVHAWGFCGGLMAHLGAPDMRAHIGFALNMPGGEVAV 283
Db 319 THFLFDVYNDIEVIVHEECIHSCEVIFIDKSVISQMYYPDMQIPILYSFWPDRIKTNL 378
QY 284 ARIDLAQASTFPKPDREPARLRADVAAGLSG---AENAKETALDHFINGR 339
Db 379 KPLDLAOSTLTFHKPSLEHPCIKLAQA---GIKNFYPYVILNMSNEIANLFLNKK 434
QY 340 IGFLDMAVVEETLAGVSTDLFGKVPDALEVLAMDLARRAAE 385
Db 435 IKYDISSIISQVLESFNSQVSENSEDLMOIILQIHSMAWDKATD 480

RESULT 12
09EX11 PRELIMINARY; PRT; 115 AA.

```

```

AC 09EX11;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE DXR reductoisomerase (fragment).
GN DXR.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CG43;
RA Lai Y.C., Chang H.Y., Peng H.L.;
RT "Identification of Klebsiella pneumoniae CG43 genes specifically
expressed in vivo.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ292312; GAC19426.1;
DR InterPro: IPR003821; DXP_reductoisomase.
DR Pfam: PF02670; DXP_reductoisom; 1.
KW isomerase.
FT NON_TER
SQ SEQUENCE 115 AA; 12174 MW; 0B5E97D06F644C8C CRC64;

Query Match 9.1%; Score 181.5; DB 2; Length 115;
Best Local Similarity 41.9%; Pred. No. 0.0002;
Matches 49; Conservative 14; Mismatches 49; Indels 5; Gaps 3;

QY 1 MRSISFGATSGESPFEDLVMRKGPEAFRTVALTGSNIRRLAEMALAEAVTAE 60
Db 1 MKQLVLSGTSIGCSTLDVY--RHNGRFSVALVACKNVDRAVEOCLETTPRYAVWD 58
QY 61 EDCJPALEAL--AGTGEVAGQAQIAE-AADRPADWTSAYGAAGLVGMRALK 114
Db 59 AQSAERLRLTRHSGRREVLSGQAAAEVAALDEVDQVMAIYGAAGLVPTLAIR 115

RESULT 13
091ZE6 PRELIMINARY; PRT; 2555 AA.
AC 091ZE6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Beta4-spectrin.
GN SPNB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21419653; PubMed=11528393;
RA Parkison N.J., Olsson C.L., Hallows J.L., McKee-Johnson J.,
RA Keogh B.P., Noben-Trauth K., Kujawa S.G., Tempel B.L.;
RT "Mutant beta-spectrin 4 causes auditory and motor neuropathies in
quivering mice.";
RL Nat. Genet. 29:61-65(2001).
DR EMBL: AY032655; AAK38731.1;
DR MCD: MGI:1890574; Spnb4.
DR InterPro: IPR001589; Actbind_actnin.
DR InterPro: IPR001715; Calponin-like.
DR InterPro: IPR001849; PH.
DR Pfam: PF00307; CH; 2.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00435; Spectrin.
DR PROSITE: PS00019; ACTININ_1; UNKNOWN_1.
DR PROSITE: PS00020; ACTININ_2; UNKNOWN_1.
DR PROSITE: PS50021; CH; 2.
DR PROSITE: PS50003; PH DOMAIN; 1.
SQ SEQUENCE 2555 AA; 288132 MW; 39FB7E410289D45 CRC64;

```


[illegible]

```

QY 129 TAGOLMRTAENGATILPVDSEHSAYFOALAGEDTACVERVIITASGSGPFRDWSLERIR 188
Db 196 LAAOTL-----ALA-----ARRALTSJMGDP--DFEIELVE 224
QY 189 --ACTVAEAOAHPPNWSMGORISIDSASMFNKALELIETREFFGFEPD-RIEAVHPOSIV 245
Db 225 PLAAPTPOMAVDPNISLVDVLAQAERAAAEAAIE---RERAAGDPDLTIOA----- 272
QY 246 HAWVFCGCGIMAHILGPDMDRHAIGFALNMP---GRGEVVARID-----LAQIA 292
Db 273 -----GYRREFQGTDRALIVGFTAPIPIRDRNQCNVAAARADANAQAARERLALTG 323
QY 293 SLTFOKPDDEERFPALRLARDVMAARGLSGAAPNNAKETALDHTAGRIGFLDMAAVEET 352
Db 324 SIRAMRDAQASIKAAEAEALVLSRTYPOA--QOAVDLAROGFEAGKFSLLDVLDQAAL 381
QY 353 LAGVTDPLFGKVPDPALEEVLMMDHLARRAAE 385
Db 382 LA--SRNDL---IAARLERALALAAERAAOE 409

```

Search completed: April 15, 2003, 13:59:51
 Job time : 93 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:58:20 ; Search time 19 Seconds
(without alignments)
1267.768 Million cell updates/sec

Title: US-09-673-198-30

Sequence: 1 MRSLSIFGATSGIGSTFDL.....MDHLARRAEAGLRQQR 394
1984

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	39.2	394	9	US-09-941-947A-8
2	778	39.2	394	10	US-09-934-903-4
3	778	39.2	394	10	US-09-934-868-64
4	741.5	37.4	477	10	US-09-923-556-6
5	741.5	37.4	477	10	US-09-987-025-2
6	737.5	37.2	477	10	US-10-047-412A-29
7	737.5	37.2	477	10	US-09-923-556-2
8	736	37.1	477	9	US-10-047-412A-10
9	628	31.7	436	9	US-09-712-363-262
10	601	30.3	392	9	US-09-738-626-5709
11	115	5.8	5215	9	US-09-860-846-2
12	115	5.8	5215	10	US-09-861-289-2
13	113.5	5.7	7257	9	US-10-014-717-5
14	106.5	5.4	609	9	US-09-308-207-67
15	106.5	5.4	1616	9	US-09-712-363-262
16	103.5	5.2	4613	9	US-09-860-846-31
17	103.5	5.2	4613	10	US-09-861-289-31
18	103.5	5.2	11877	9	US-09-860-846-6
19	103.5	5.2	11877	10	US-09-861-289-6

20	102	5.1	1346	9	US-09-860-846-37	Sequence 37, Appl
21	102	5.1	1346	10	US-09-861-289-37	Sequence 37, Appl
22	101	5.1	538	9	US-09-738-626-4165	Sequence 4165, Ap
23	100.5	5.1	523	10	US-09-815-242-13254	Sequence 13254, A
24	100.5	5.1	535	9	US-09-712-363-228	Sequence 228, App
25	99.5	5.0	1407	10	US-09-815-242-10439	Sequence 10439, A
26	99.5	5.0	1407	10	US-09-815-242-14042	Sequence 14042, A
27	98.5	5.0	419	9	US-09-712-363-289	Sequence 289, App
28	98	4.9	702	10	US-09-815-242-4975	Sequence 4975, Ap
29	98	4.9	704	10	US-09-815-242-10882	Sequence 10882, A
30	97.5	4.9	3739	9	US-09-860-846-33	Sequence 33, Appl
31	97.5	4.9	3739	10	US-09-861-289-33	Sequence 33, Appl
32	96.5	4.9	523	10	US-09-814-041A-2	Sequence 2, Appl
33	96.5	4.9	888	9	US-09-738-626-5302	Sequence 5302, Ap
34	95.5	4.8	527	9	US-09-738-626-5154	Sequence 5154, Ap
35	95.5	4.8	3519	9	US-09-808-880-4	Sequence 4, Appl
36	94.5	4.8	1333	9	US-09-808-880-3	Sequence 4048, Ap
37	94.5	4.8	3816	9	US-09-808-880-2	Sequence 3, Appl
38	94.5	4.8	4150	9	US-09-808-880-2	Sequence 2, Appl
39	93	4.7	466	10	US-09-815-242-13525	Sequence 13525, A
40	92.5	4.7	1562	9	US-09-860-846-35	Sequence 35, Appl
41	92.5	4.7	1562	10	US-09-861-289-35	Sequence 35, Appl
42	92.5	4.7	7968	9	US-10-077-130-5	Sequence 5, Appl
43	92	4.6	513	10	US-09-833-745-63	Sequence 63, Appl
44	92	4.6	567	9	US-09-957-960-2	Sequence 2, Appl
45	91.5	4.6	1289	9	US-09-712-363-259	Sequence 259, App

ALIGNMENTS

RESULT 1
US-09-941-947A-8
Sequence 8, Application US/09941947A
Publication No. US2003003528A1
GENERAL INFORMATION:
APPLICANT: Bzostowicz, Patricia C.
APPLICANT: Cheng, Qiong
APPLICANT: Dicosimo, Deana J.
APPLICANT: Koffas, Mattheos
APPLICANT: Miller, Edward S. Jr.
APPLICANT: Odom, J. Martin
APPLICANT: Picataggio, Steve
APPLICANT: Rouverelle, Pierre E.
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
FILE REFERENCE: C11903 US NA
CURRENT APPLICATION NUMBER: US/09/941,947A
CURRENT FILING DATE: 2001-09-01
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Microsoft Office 97
SEQ ID NO 8
LENGTH: 394
TYPE: PRT
ORGANISM: Methylobionas 16a
US-09-941-947A-8
Query Match 39.2%; Score 778; DB 9; Length 394;
Best Local Similarity 41.6%; Pred. No. 5e-60;
Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps. 5;
1 MRSLSIFGATSGIGSTFDLVMKGPDEAFRIYALGGRNIRLAEMARLKLAVTAH 60
1 MKGICILGATSGIGSTFDLVARRH--PDKYQVALPLANGSIDALYECIAHHEVAVVWA 58.
DB
61 EDCLPLALREALT---GFEVAGCAQAIAD--RADMTMSAIVGAGIVPMRLKKG 116
DB
59 ESKVAEFKRIASPVADIKVLSGSEPLQOVATLENDYVAIVGAGILPLLAARKG 118
OY 117 RTIALANKESLVTAGOLMLMRTAENGATILPVSEHSAAVFOAL-----AGEDTACVERVI 171

```
Db 119 KTVLLANKREALVMSGQIFMCAVSDGAVLLPIDSEHNAIFQCMRGYTRGHAKQARRIL 178
Qy 172 ITASGCPFRDMSLEIRACTVAEAQAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPDOAVAHKMDGKRISVDSATMNMKGLELEACLLFNME 238
Qy 232 PDRIEAVVHPOSIVAHAVGFCDDGLMAHLGSPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNDMHTPIAHAAWMPERDGSVAPLDIEFV 298
Qy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAFAFNKAEIALDHFYAGHIGFLDMAAYVEE 351
Db 299 GHMDEKPDLRFPCLRLAYEAIKSGIMPTVLNANAEIAVEAFINEEVKFTDIAVIER 358
Qy 352 TLGAVSTDPLFGKVPDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387
```

RESULT 2

```
US-09-934-903-4
; Sequence 4, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Schenzle, Andreas J.
; APPLICANT: No. US20020102690A1ton, Kelley C.
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Rouviere, Pierre
; APPLICANT: Picataggio, Stephen
; APPLICANT: Cheng, Qiong
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
; FILE REFERENCE: C11646 US NA
; CURRENT APPLICATION NUMBER: US/09/934,903
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,907
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 394
; TYPE: PR
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: Amino acid sequences encoded by ORF2
US-09-934-903-4
```

Query Match 39.2%; Score 778; DB 10; Length 394;

Best Local Similarity 41.6%; Pred. No. 5e-60; Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;

```
Qy 1 MRSLSIFGATSGESTFDLVKRGKGPFAFTVALTGGRNIRLAEMARALKAEIAYTAH 60
Db 1 MKGCIILGATSGISVSTLDVVAHR--PDKYQVVALTANGNDIALYEQCLAHHPREYAVVM 58
Qy 61 EDCIPLARELAGT--GTEVAGGAQAIAEAAD--RPADMTSAIYGAAGLVGRALKHG 116
Db 59 ESKAEEFKQRIASPVADIKVLSSGEALQOYATLENDVTVAALYGAAGLLPTLAAAG 118
Qy 117 RTLLANKESLVTAQGLMRTAENGATILPVSEHSAVFOAL-----AGEDTACVERVI 171
Db 119 KTVLLANKREALVMSGQIFMCAVSDGAVLLPIDSEHNAIFQCMRGYTRGHAKQARRIL 178
Qy 172 ITASGCPFRDMSLEIRACTVAEAQAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPDOAVAHKMDGKRISVDSATMNMKGLELEACLLFNME 238
Qy 232 PDRIEAVVHPOSIVAHAVGFCDDGLMAHLGSPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNDMHTPIAHAAWMPERDGSVAPLDIEFV 298
```

```
Qy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAFAFNKAEIALDHFYAGHIGFLDMAAYVEE 351
Db 299 GHMDEKPDLRFPCLRLAYEAIKSGIMPTVLNANAEIAVEAFINEEVKFTDIAVIER 358
Qy 352 TLGAVSTDPLFGKVPDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387
```

RESULT 3

```
US-09-934-868-64
; Sequence 64, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 64
; LENGTH: 394
; TYPE: PR
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: Amino acid sequences encoded by DXR
US-09-934-868-64
```

Query Match 39.2%; Score 778; DB 10; Length 394;

Best Local Similarity 41.6%; Pred. No. 5e-60; Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;

```
Qy 1 MRSLSIFGATSGESTFDLVKRGKGPFAFTVALTGGRNIRLAEMARALKAEIAYTAH 60
Db 1 MKGCIILGATSGISVSTLDVVAHR--PDKYQVVALTANGNDIALYEQCLAHHPREYAVVM 58
Qy 61 EDCIPLARELAGT--GTEVAGGAQAIAEAAD--RPADMTSAIYGAAGLVGRALKHG 116
Db 59 ESKAEEFKQRIASPVADIKVLSSGEALQOYATLENDVTVAALYGAAGLLPTLAAAG 118
Qy 117 RTLLANKESLVTAQGLMRTAENGATILPVSEHSAVFOAL-----AGEDTACVERVI 171
Db 119 KTVLLANKREALVMSGQIFMCAVSDGAVLLPIDSEHNAIFQCMRGYTRGHAKQARRIL 178
Qy 172 ITASGCPFRDMSLEIRACTVAEAQAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPDOAVAHKMDGKRISVDSATMNMKGLELEACLLFNME 238
Qy 232 PDRIEAVVHPOSIVAHAVGFCDDGLMAHLGSPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNDMHTPIAHAAWMPERDGSVAPLDIEFV 298
Qy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAFAFNKAEIALDHFYAGHIGFLDMAAYVEE 351
Db 299 GHMDEKPDLRFPCLRLAYEAIKSGIMPTVLNANAEIAVEAFINEEVKFTDIAVIER 358
Qy 352 TLGAVSTDPLFGKVPDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387
```

RESULT 4

```
US-09-923-556-6
; Sequence 6, Application US/09923556
; Patent No. US20020058321A1
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
```

RESULT 5
US-09-987-025-2
Sequence 2, Application US/09987025
Patent No. US20020108148A1
GENERAL INFORMATION:
APPLICANT: Boronat, Albert
APPLICANT: Campos, Narciso
APPLICANT: Kishore, Ganesh M.
TITLE OF INVENTION: Nucleic Acid Sequences
Involved In Isoprenoid Synthesis
TITLE OF INVENTION: Isoprenoid Synthesis
FILE REFERENCE: 17142/02/US
CURRENT APPLICATION NUMBER: US/09/987,025
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: 09/549,787
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/129,899
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 60/146,461
PRIOR FILING DATE: 1999-07-30
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0

```

1 RESULT 6
2 US-10-047-412A-29
3 Sequence 29, Application US/10047412A
4 Publication No. US20020197696A1
5 GENERAL INFORMATION:
6 APPLICANT: Levin, Joshua Z.
7 APPLICANT: Budziszewski, Gregory J.
8 APPLICANT: Potter, Sharon L.
9 APPLICANT: Weiglich, Lynette M.
10 TITLE OF INVENTION: Herbicide Target Genes and Methods
11 FILE REFERENCE: PP/5-30780DIY
12 CURRENT APPLICATION NUMBER: US/10/047,412A
13 CURRENT FILING DATE: 2002-04-11
14 NUMBER OF SEQ ID NOS: 29
15 SOFTWARE: PatentIn Ver. 2.1
16 SEQ ID NO 29
17 LENGTH: 477
18 TYPE: PRT
19 ORGANISM: Arabidopsis thaliana
20 FEATURE:
21 NAME/KEY: SITE
22 LOCATION: (39)
23 OTHER INFORMATION: Xaa = Asp or Asn
24 FEATURE:
25 NAME/KEY: SITE
26 LOCATION: (155)
27 OTHER INFORMATION: Xaa = Pro or Leu
28 FEATURE:
29 NAME/KEY: SITE
30 LOCATION: (162)
31 OTHER INFORMATION: Xaa = Leu or Gln
32 FEATURE:
33 NAME/KEY: SITE
34 LOCATION: (187)

```

OTHER INFORMATION: Xaa - Lys or Gln
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (465)
 OTHER INFORMATION: Xaa - Asp or Asn
 US-10-047-412A-29

Query Match 37.2%; Score 737.5; DB 9; Length 477;
 Best Local Similarity 41.1%; Pred. No. 2.2e-56;
 Matches 161; Conservative 71; Mismatches 141; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGTSTPDLVNRKGGPEAFRTVALTGNNIRRLAEMARALAEAVTAHE 61
 DB 81 KRISIVGTSIGTQTLDIYAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAAVNE 138
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP--ADMTMSAIVGAGLVPGMRALKHGT 118
 DB 139 SLINELKALADLDYKKEIIIEGEGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMRTAENGATILPVDESHSAVFOALAGEBTACVERVIITASGP 178
 DB 199 IALANKETLLIAGGFVLPRLANKHNKILPADSEHSAIFQCIOGLPEGALRKIIILTASGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMNSGCRISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNMNMNGKITVDSCITLFNKGLEVIEAHYLFGEAYDDIEIV 318
 QY 239 VHPQSIYAMVGFCDGGLMAHLGPADMRAHAGFALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETQDSSVLAQLGMPDMRLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAFAFNAAKELADHFTAGRIGFLDMAAIVEETL-- 353
 DB 379 FKRPDNNVYPPSMDLAAAGAGGTMTGVLSAANEKAAVEFIDEKISTYLDIKFVVELTCDK 438
 QY 354 --AGVSTDPFLGKVPDALEEVLAHDLARRAA 383
 DB 439 HRNELVTSPP-----SLEIVHYDLMAREYA 463

RESULT 7
 US-09-923-556-2
 Sequence 2, Application US/09923556
 Patent No. US20020058321A1
 GENERAL INFORMATION:
 APPLICANT: Busch, Marco
 APPLICANT: Hain, Rudiger
 APPLICANT: Martin, William
 APPLICANT: Tietjen, Klaus
 APPLICANT: Klotl, Andreas
 TITLE OF INVENTION: Method of determining the activity of
 TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
 FILE REFERENCE: 2020US
 CURRENT FILING DATE: 2001-08-07
 PRIOR APPLICATION NUMBER: US/09/923,556
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 2
 LENGTH: 477
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-09-923-556-2

Query Match 37.2%; Score 737.5; DB 10; Length 477;
 Best Local Similarity 41.1%; Pred. No. 2.2e-56;
 Matches 161; Conservative 71; Mismatches 141; Indels 19; Gaps 6;
 QY 2 RSLIFGATGSGTSTPDLVNRKGGPEAFRTVALTGNNIRRLAEMARALAEAVTAHE 61
 DB 81 KRISIVGTSIGTQTLDIYAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAAVNE 138

QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP--ADMTMSAIVGAGLVPGMRALKHGT 118
 DB 139 SLINELKALADLDYKKEIIIEGEGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMRTAENGATILPVDESHSAVFOALAGEBTACVERVIITASGP 178
 DB 199 IALANKETLLIAGGFVLPRLANKHNKILPADSEHSAIFQCIOGLPEGALRKIIILTASGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMNSGCRISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNMNMNGKITVDSCITLFNKGLEVIEAHYLFGEAYDDIEIV 318
 QY 239 VHPQSIYAMVGFCDGGLMAHLGPADMRAHAGFALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETQDSSVLAQLGMPDMRLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAFAFNAAKELADHFTAGRIGFLDMAAIVEETL-- 353
 DB 379 FKRPDNNVYPPSMDLAAAGAGGTMTGVLSAANEKAAVEFIDEKISTYLDIKFVVELTCDK 438
 QY 354 --AGVSTDPFLGKVPDALEEVLAHDLARRAA 383
 DB 439 HRNELVTSPP-----SLEIVHYDLMAREYA 463

RESULT 8
 US-10-047-412A-10
 Sequence 10, Application US/10047412A
 Publication No. US20020197696A1
 GENERAL INFORMATION:
 APPLICANT: Levin, Joshua Z.
 APPLICANT: Budziszewski, Gregory J.
 APPLICANT: Potler, Sharon L.
 APPLICANT: Megrich, Lynette M.
 TITLE OF INVENTION: Herbicide Target Genes and Methods
 FILE REFERENCE: PB/5-30780DIV
 CURRENT FILING DATE: 2002-04-11
 CURRENT APPLICATION NUMBER: US/10/047,412A
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 10
 LENGTH: 477
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-10-047-412A-10

Query Match 37.1%; Score 736; DB 9; Length 477;
 Best Local Similarity 40.5%; Pred. No. 3e-56;
 Matches 163; Conservative 74; Mismatches 143; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGTSTPDLVNRKGGPEAFRTVALTGNNIRRLAEMARALAEAVTAHE 61
 DB 81 KRISIVGTSIGTQTLDIYAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAAVNE 138
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEADRP--ADMTMSAIVGAGLVPGMRALKHGT 118
 DB 139 SLINELKALADLDYKKEIIIEGEGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMRTAENGATILPVDESHSAVFOALAGEBTACVERVIITASGP 178
 DB 199 IALANKETLLIAGGFVLPRLANKHNKILPADSEHSAIFQCIOGLPEGALRKIIILTASGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMNSGCRISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNMNMNGKITVDSCITLFNKGLEVIEAHYLFGEAYDDIEIV 318
 QY 239 VHPQSIYAMVGFCDGGLMAHLGPADMRAHAGFALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETQDSSVLAQLGMPDMRLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAFAFNAAKELADHFTAGRIGFLDMAAIVEETL-- 353

Db 379 FKPDNVKYPMDLAYAAGAGTMTGVLSPAANEKAVEMIDEKISYLDIFKVELTCDK 438
QY 354 --AGVSTDPLEFGKVPDALEEVLAHDLARRAAE---AAGLR 390
Db 439 HNEELVTSP-----SLEIYHYDLMAREYADVOLSSGAR 473

RESULT 9

US-09-712-363-258
; Sequence 258 Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258
; LENGTH: 436
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-258

Query Match 31.7%; Score 628; DB 9; Length 436;

Best Local Similarity 39.2%; Pred. No. 6,9e-47;

Matches 155; Conservative 48; Mismatches 152; Indels 40; Gaps 7;

QY 6 IFGATGSIESTFEDLMRRGGEPAFTVALTGGRNIRLAEMARALKAEIATVAHEDCLP 65
Db 40 VLASTSISTGTOALQVY--ADNPDREFVVGILAG-----GAHLDTLL 78
QY 66 ALREALAGTGEVAG--GAQAI-----AEADR-----PADMTSAIYGAAGLVPGMR 111
Db 79 RQRAQGVNINAVADHHAQVRGDIPIHGSDATRLVEQTEADYVNLALVGALGRLTILA 138
QY 112 ALKHGTTALANKESLVYTGOLIMRTAENGATILPYDSEHSAVFQALAGEPTACVERVY 171
Db 139 AKTGRLRLANKESLVAGGSILVRAARP--GQIVPDESHSALAOCLRGSTPDEYAKLV 196
QY 172 ITASGGEFDMWSLERIACVTAEAQAHPMWSMGORISIDSASMFNKALEIETREFGE 231
Db 197 LTRSGGPFEGMSADIEHVTPEQAGNHPYWSGPMNTLASASLVNKGLEYETHLFGIP 256
QY 232 PRIEAVVHPOSIVHAMVFCDDGLMAHLGPDADMRHAIGFALNMPGREGVPAVIDLAQI 291
Db 257 YRIDIVVHPOSIIHSMVTFIDGSTIAQASPPMKLPISLALGMPRRVSGAAACPFHNA 316
QY 292 ASITFGKPEDEERFRLRLADVMAARGLSGAFFNAKETALDHFITAGTIGFLDMAVVEE 351
Db 317 SSWEFEPDLTDVFPVAVELARQAGVAGGCMYAVYNAANEEMAAFLAGRIGFPAIVGIAD 376

QY 352 TLAGVSTDPLEFGKVPDALEEVLAHDLARRAAEA 386
Db 377 VLHAADQ---WAVEPATVDVLDLQAKRARRAORA 408

RESULT 10

US-09-738-626-5709
; Sequence 5709 Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5709
; LENGTH: 392
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5709

Query Match 30.3%; Score 601; DB 9; Length 392;

Best Local Similarity 37.4%; Pred. No. 1.3e-44;

Matches 146; Conservative 67; Mismatches 155; Indels 22; Gaps 10;

QY 2 RLSITFGATGSIESTFEDLMRRGGEPAFTVAL--TGGRNIRLAEMAR--ALKAEIATV 58
Db 6 KTIILIGSTGISTGTOALDVI--ADNSDKREFVVGILAGSOPDLVISOAQOGLIAAKRV 63
QY 59 AHEDCLPALREALAG--TGTEVAGGAQAIAEADRPAADMTSAIYGAAGLVPGMRALKH 115
Db 64 ADQAAAVASKALGELISTD---AAKILVETT--KADIVNALVIGSLGATLATLTLS 118
QY 116 GRTLALANKESLVYTAGOLIMRTAENGATILPYDSEHSAVFQALAGEPTACVERVYIT 175
Db 119 GAHLALANKESLVAGGEFV--TSKARLGOIIPDSEHSAVMAOCLRGSTPDEYARIYLTAS 176
QY 176 GGEFDMWSLERIACVTAEAQAHPMWSMGORISIDSASMFNKALEIETREFGEPPDR 235
Db 177 GGEFDMWSLERIACVTAEAQAHPMWSMGORISIDSASMFNKALEIETREFGEPPDR 236
QY 236 EAVVHPOSIVHAMVFCDDGLMAHLGPDADMRHAIGFALNMPGREGVPAVIDLAQIAS 293
Db 237 DVYVHPOSIIHSMVTFIDGSTIAQASPPMKLPIALALDMPNR--VPKQPALDFTAHT 294
QY 294 LTRQKDEERFRLRLADVMAARGLSGAFFNAKETALDHFITAGTIGFLDMAVVEEL 353
Db 295 WAEFEPDADAFAVOLARHAKOKGTPYAVYNAANEEMAAEAFIRGRIRKPOIYDVDEV 354
QY 354 AGVSTDPLEFGKVPDALEEVLAHDLARRAA 383
Db 355 QGASQ---FAGVASHVDILATESEARARA 381

RESULT 11

US-09-860-846-2
; Sequence 2, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-2

Query Match 5.8%; Score 115; DB 9; Length 5215;
Best Local Similarity 20.2%; Pred. No. 1.1;
Matches 108; Conservative 52; Mismatches 162; Indels 212; Gaps 21;

QY 28 EAFRTVALTGGRNIRRLAEMARALKAEIA-----VTANEDCLPALREALAGTGEVAG 80
DB 1509 QAVRLRTFLDGRPDVTVAADGRSLAARTAFEHKALTTATFDELLAGIDLAL-GRGEQATG 1567
QY 81 -----GAQAIA-----GLVP-----GMRALKH-GRITLANK 125
DB 1568 LVTGEPARAGRTAFLTGCGAQRVAMGEELRAHPVFAALDTVYALDRHLDRPLREIV 1627
QY 102 GAA-----GLVP-----GMRALKH-GRITLANK 125
DB 1528 AAGEELDTAYTOPALFAFEVALFRLEHGLVDPDLTGHSVGEIAAHVAGVSLDDAA 1687
QY 126 SLVTAGOLMRTAOENCATILPVDSHSAVFOALAGEDTACVEVITTAGSGCP----- 179
DB 1688 RLVTARGLMOSAREGA-MIAVAGAEVAVESLKYE---GRVAAVAVNGPAAVAVVSG 1742
QY 180 -RDMSLERIRACTVAEAOAHNMS-MGQRI-----SIDASMFNKALIELTREFEGF 230
DB 1743 DAD-AAEIRIV-----MAGRGRTRLRVSHAFSPHMDVLDLFLVAEGILTF 1791
QY 231 EPRDIEAVHPOSIVHAMV-----GFCDDGLMAHLGP 262
DB 1792 EEPRIPIV---STVTGALVTSGETLSPAYWVDQIRRPVRFDAVRTLAADATVLEIGP 1848
QY 263 ADMHAIGFALNMPGRGEVPAVIDLAQIASLTQKDEE----- 302
DB 1849 DAVTLAAEELALPGT-DAPDAR-DVTIVPRLLRAGREPPELTAAGLATAHVHGAPLDRAS 1906
QY 303 -----RPPALRLANDVMAARG--LSGA 322
DB 1907 FFPDGRRTDPTVAFRRREHWLTPPEARNDARALGFDPARHPDLTTTVEVAGDGVLLTGR 1966
QY 323 AFNAKEIALDHFIFAGRI-----GFLDMAVVEETLAGVSTDPLEFGVPDALEE 371
DB 1967 LSLTDQWMLADHMVNGAVLLPATATFELALAAAGDHGAVRVEELTLAPLPLPE 2020

RESULT 12
US-09-861-289-2
; Sequence 2, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-2

Query Match 5.8%; Score 115; DB 10; Length 5215;
Best Local Similarity 20.2%; Pred. No. 1.1;
Matches 108; Conservative 52; Mismatches 162; Indels 212; Gaps 21;

QY 28 EAFRTVALTGGRNIRRLAEMARALKAEIA-----VTANEDCLPALREALAGTGEVAG 80
DB 1509 QAVRLRTFLDGRPDVTVAADGRSLAARTAFEHKALTTATFDELLAGIDLAL-GRGEQATG 1567
QY 81 -----GAQAIA-----GLVP-----GMRALKH-GRITLANK 125
DB 1628 AAGEELDTAYTOPALFAFEVALFRLEHGLVDPDLTGHSVGEIAAHVAGVSLDDAA 1687
QY 126 SLVTAGOLMRTAOENCATILPVDSHSAVFOALAGEDTACVEVITTAGSGCP----- 179
DB 1688 RLVTARGLMOSAREGA-MIAVAGAEVAVESLKYE---GRVAAVAVNGPAAVAVVSG 1742
QY 180 -RDMSLERIRACTVAEAOAHNMS-MGQRI-----SIDASMFNKALIELTREFEGF 230
DB 1743 DAD-AAEIRIV-----MAGRGRTRLRVSHAFSPHMDVLDLFLVAEGILTF 1791
QY 231 EPRDIEAVHPOSIVHAMV-----GFCDDGLMAHLGP 262
DB 1792 EEPRIPIV---STVTGALVTSGETLSPAYWVDQIRRPVRFDAVRTLAADATVLEIGP 1848
QY 263 ADMHAIGFALNMPGRGEVPAVIDLAQIASLTQKDEE----- 302
DB 1849 DAVTLAAEELALPGT-DAPDAR-DVTIVPRLLRAGREPPELTAAGLATAHVHGAPLDRAS 1906
QY 303 -----RPPALRLANDVMAARG--LSGA 322
DB 1907 FFPDGRRTDPTVAFRRREHWLTPPEARNDARALGFDPARHPDLTTTVEVAGDGVLLTGR 1966
QY 323 AFNAKEIALDHFIFAGRI-----GFLDMAVVEETLAGVSTDPLEFGVPDALEE 371
DB 1967 LSLTDQWMLADHMVNGAVLLPATATFELALAAAGDHGAVRVEELTLAPLPLPE 2020

RESULT 13
US-10-014-717-5
; Sequence 5, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0


```

; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
;
US-10-014-717-5

```

Query Match	5.7%;	Score 113.5;	DB 9;	Length 7257;
Best Local Similarity	24.0%;	Pred. NO. 2.4;		
Matches 95;	Conservative 42;	Mismatches 136;	Indels 123;	Gaps 19

OY	4	LSIFGAAGSICESEFDDVMKGGPEAFPRVATLGGNINIRLEMMRAKKAELA-----VT	58
Dd	5510	VSSPGVS-----TNAVHVELEAPAAFPAPAARSAELFVLSKSAALDQAQARSANHV	5565
OY	59	AHEBC-IPALREALAGTEV-----AGCAOAIIEAADRPADWTMSAIVGMAALVGCM	110
Dd	5566	AHPELGIGDLAFSLATRSMYTRLAVAATSRRLASAID-----TAAGQAPPA	5615
OY	111	RALKHGTTLLANKESLVTAGQLMRTAOBNGATTLPVDSESAVFQALAGED-----	163
Dd	5616	AARGHAFT-GSAPKVVFVFGQ-----GSOWLMGMOKLISEEPVRDAL	5658
OY	164	TACVERIITASGPFFWDMLERACTYAEOAHNMMSGORISIDSMSMNKLJELLE	223
Dd	5659	SAC-DRIAQEAG-----WSL-----LAELLADETTSQLRIDVOPALPIEVALSA	5705
OY	224	TREFFGEPPRIEAUVHPOSIVHANVGFCDGLMAHLGPDNRHAI GFALNMPGRCEVPV	283
Dd	5706	LMRSMGVCPAY-----VGISM-----GEVAA	5722
OY	284	ARIIDLQIASLTQKPDERFPALRLARDYMAARGLSGAAFNAKEIALDHFIAGRIGFL	343
Dd	5728	AHV--AGALSU-----EDAVAI-ICRSILLRISGGEMAVELLSLAEALALLGYE	5777
OY	344	DMAVVEETLAGYSTDP-----LGCKPDALEEVLAAM	375
Dd	5778	DRUSV-----AVSNSPRSTVLAGE-PALAELVELI	5806

RESULT 14
US-09-308-207-67
Sequence 67, Application US/09308207
Publication No. US20030022323a1
GENERAL INFORMATION:
APPLICANT: MARIA DIAZ-TORRES ET AL.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
PRODUCTION OF 1,3 PROPANEDIOL
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
1870 South Winton road
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,207
FILING DATE: 13-May-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-864-7620
 TELEFAX: 650-845-6504
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 67:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 609 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: NO. US2003002323A1e
 SEQUENCE DESCRIPTION: SEQ ID NO: 67
 US-09-308-207-67

Query Match	5.4%	Score 106.5	DB 9	Length 609
Best Local Similarity	23.5%	Pred. No. 0.32		
Matches 96	Conservative 48	Mismatches 116	Indels 149	Gaps 22

QY	27	PERPVIAITGCRNIRLRAEAKRL---	KAEIAVTHED---	CLALREALAGT---	75
Db	214	PYGATATFEDLMAEESQNVPMARL	IGNRSAVVVKTSQGVKARATPAGN	ULLIAQGRSV	273
QY	76	-TEVAGQAQIAIEAADRPADMTSAI	VGAGL--VPGRALKHGRITLALANK	ESLVTAGOL	133
QY	274	QVDAVAAAGEALIMKAVD--	CGCKLDNVAGEAGNTNGM--	LEHVRQ-----	TIANEL
Db	134	LMRTAENGATILPVDSEBSAVFOAL	AGEDACVERVITPASGPPFDMSL	ERACTVA	193
QY	320	TNRPAQE-----	IRIDDLAVDTA--	PPVSTGTGLAGEEFSLEQ-----	355
QY	194	EAQAHPNMSQGRISIDSASFENKAL	ELIETREFGEPRIEAVNPQSI	VHANMGCD	255
Db	356	-----AVGIASM--	-----	VKSRLQALLARELEHL--	381
QY	254	GLMAHLGPADMRHAIGFALNMP	PERGEVVARIDL-----	AQIASLTFOKPEDEREPAL	307
Db	382	QIAVQVQAGAAEALIGALTPTTTR-	PLAIIIDLGAGSTDA	SIIN-----	425
QY	308	RLADVMAARGLSGA----	AFNAKKEIAL-DHETAGRIGFLMA	AAVEETL-----	353
Db	426	-AAGELISATHLAGAGDWVTMII	IRELIGLEDRERYLAEEIKKY	PLAKV--ESLFLHRHEDGS	481
QY	354	-----AGVSTD--	PLGKVPDAALEEVLA	MDHLARRAE	384
Db	482	VOFPSPALPTVFEARVCVKPEL	VPGLGDP--LEKVRAT---	RSRAK	524

```

RESULT 15
US-09-712-363-262
; Sequence 262, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206,
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14

```



```

; FEATURE:
; OTHER INFORMATION: ORF2
US-09-934-903-3
Query Match 19.0%: Score 224.2; DB 10; Length 1182;
Best Local Similarity 56.5%: Pred. No. 2.5e-43;
Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

Db 403 ATCTTCATGACAGCCGCTCAGCGATTCGCCGCTGTGTGCTGCGATAGACAGCGAC 462
QY 457 TCCGGGCTTTTCAGGCGCTCGGGGCGAGACACGGCTC-----CTC 501
Db 463 AACGCACTCTTTCAGTGCATCCCGCGGTTATACGCCAGCCATACAGCAACAGCGC 522
QY 502 GAGCGGCTCATATCAGCGCGTCCGCGCGGCGCTTCCGCGATGAGCGCTCGAGCGCATC 561
Db 523 CGCGGCATTTTATGACCGGCTTCCGGTGGCCATTTGAGAGAGCGCCATAGAAACGTT 582
QY 562 CGCGCTGACACGCTGCGGAGCGGAGCCCATCCCACTGGTCCATGCGCCAGCGATC 621
Db 583 TCCAGCGCTACGCGGATCAGCGCGCTTCCCATCTTAATGAGCATGCGCGCAAGATT 642
QY 622 TCCATGACAGCGCTCATGTTCAACAGGCGCTCGAGCTGATGAGAGCGCGCAATTC 681
Db 643 TCGGTCGATTCGCGCCACCATATGACAAAGGTCTCGACTGATGAGAGCTGCTTGTG 702
QY 682 TTGCGCTTCAGACCGGAGCGGATGAGCGGCTGCTCATCCGCAATCCATGCTCATCG 741
Db 703 TTCAACATGAGAGCCGACAGATTGAAGTGTATTCGATCCGAGCATCATTCATTG 762
QY 742 ATGCTGGCTTCTGCGAGGCGGCTGATGCGCCATCTCGGCGCCGCGCATGCGCAC 801
Db 763 ATGCTGACTATGCTGATGCTGCTGCTTTGGCGAGATGGGTAAATCCGACATGCGCAG 822
QY 802 GCCATCGGATTCGCGCTGACACTGCGCGGCTGCGCGGAGTGCCTGCGCGGATTCAC 861
Db 823 CCGATAGCGGACGGATGCGCTGCGCGGAGACGCTTTGACTGTGGTGGCGCGCTGAT 882
QY 862 CTCGACAGATTTGAGGCTCAGCTTCCAGAGCCTGACAGAGAACGTTTCCGCGCTG 921
Db 883 ATTTTCGAGTAGGCGACATGATTTTCGAAACCCGACTTGAAACGGTTTCTGTCTG 942
QY 922 AGGCTGGCGGAGAGCTATGCGGCGGCGGCGCTGTCGCGCGCGCTTAAAGCGCGCC 981
Db 943 AGATTGGCTTATGAAAGCATATAGCTGTGTGAATATATGCCAAGCGTATTAAGCGCAC 1002
QY 982 AAGAGATCGCGCTGATATTCATGCGCCGAGCGCATGCGGTTTCTGAGATGCGCGC 1041
Db 1003 AATGAATTCGTCGCGAAGCGTTTAAATGAAGAAGTCAATTCACATGACATGCGGTC 1062
QY 1042 GTGCTCGAG 1050
Db 1063 ATCATCGAG 1071

RESULT 2
US-09-934-903-3
; Sequence 3, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Mattheos
; APPLICANT: Schenkle, Andreas J.
; APPLICANT: No. US20020102690A1ton, Kelley C.
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Rouviere, Pierre
; APPLICANT: Plicataggio, Stephen
; APPLICANT: Cheng, Olong
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
; FILE REFERENCE: C11646 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 903
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,907
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Methylobionas 16a

```

```

; FEATURE:
; OTHER INFORMATION: ORF2
US-09-934-903-3
Query Match 19.0%: Score 224.2; DB 10; Length 1182;
Best Local Similarity 56.5%: Pred. No. 2.5e-43;
Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

Db 277 GCCGCTGACAGCTGATGCGGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 336
QY 283 GTGATAGCGGATGATGCGGCTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 342
Db 337 CTGAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 396
QY 343 GCCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 402
Db 397 CTCTGATGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 456
QY 403 ATCTTCATGACAGCGCTCATGATTCGCGCGCTGTGTGCTGCGGATAGACAGCGAGC 462
Db 457 TCCGCGCTTTCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 501
QY 463 AACGCACTCTTTCAGTGCATCCCGCGGTTATACGCCAGCCATACAGCAACAGCGC 522
Db 502 GAGCGGCTCATATCAGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 561
QY 523 CGCGCATTTTATGACCGCTTCCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 582
Db 562 CGCGCGCTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 621
QY 583 TCCAGCTGACGCGCGGATCAGCGCGCTTCCCATCTTAATGAGCATGCGCGCAAGATT 642
Db 622 TCCATGACAGCGCTCATGTTCAACAGGCGCTGCGAGCTGATGAGAGCGCGCAATTC 681
QY 643 TCGGTCGATTCGCGCCACCATATGACAAAGGTCTCGACTGATGAGAGCTGCTTGTG 702
Db 682 TTGCGCTTCAGACCGGAGCGGATGAGCGGCTGCTCATCCGCAATCCATGCTCATGCG 741
QY 703 TTCAACATGAGAGCCGACAGATTGAAGTGTATTCGAGAGCATATTCATTCG 762
Db 742 ATGCTGGCTTCTGCGAGGCGGCGCTGATGCGCCATCTCGGCGCGCGCGCGCATGCGCAC 801
QY 763 ATGCTGACTATGCTGATGCTGCTTGGCGAGATGGGTAAATCCGACATGCGCAG 822
Db 802 GCCATCGGATTCGCGCTGACACTGCGCGGCTGCGCGGAGTGCCTGCGCGGATTCAC 861
QY 823 CCGATAGCGGACGGATGCGCTGCGCGGAGACGCTTTGACTGTGGTGGCGCGCTGAT 882
Db 862 CTCGACAGATTTGAGGCTCAGCTTCCAGAGCCTGACAGAGAACGTTTCCGCGCTG 921
QY 883 ATTTTCGAGTAGGCGACATGATTTTCGAAACCCGACTTGAAACGGTTTCTGTCTG 942
Db 922 AGGCTGGCGGAGAGCTATGCGGCGGCGGCGCTGTCGCGCGCGCTTAAAGCGCGCC 981
QY 943 AGATTGGCTTATGAAAGCATATAGCTGTGTGAATATATGCCAAGCGTATTAAGCGCAC 1002
Db 982 AAGAGATCGCGCTGATATTCATGCGCCGAGCGCATGCGGTTTCTGAGATGCGCGC 1041
QY 1003 AATGAATTCGTCGCGAAGCGTTTAAATGAAGAAGTCAATTCACATGACATGCGGTC 1062
QY 1042 GTGCTCGAG 1050
Db 1063 ATCATCGAG 1071

RESULT 3
US-09-934-868-63
; Sequence 63, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenkle, Andreas J

```

```

: TITLE OR INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN
: FILE REFERENCE: CIL1596 US NA
:
: CURRENT APPLICATION NUMBER: US/09/934, 868
: CURRENT FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/7229, 858
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 81
: SOFTWARE: Microsoft Office 97
:
: SEQ ID NO 63
:
: LENGTH: 1182
:
: TYPE: DNA
:
: ORGANISM: Methylomonas 16a
: FEATURE:
: OTHER INFORMATION: DMR
:
: US-09-934-868-63

```

[illegible]

```

1      |||||
Db      1063  ATCATCGAG 1071

RESULT 4
US-09-938-956-5
; Sequence 5, Application US/09938956
; Patent No. US20020142408A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Siqun
; APPLICANT: Dicosimo, Deana J.
; APPLICANT: Kofas, Matheos
; APPLICANT: Odom, J. Martin
; TITLE OF INVENTION: Production of Monoterpene
; FILE REFERENCE: C11809 US NA
; CURRENT APPLICATION NUMBER: US/09/938,956
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-0
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 11575
; TYPE: DNA
; ORGANISM: Plasmid
US-09-938-956-5

```

Query Match	Similarity	19.0%	Score 224.2	DB 10	Length 11575
Best Local	Similarity	16.5%	Pred. No. 2e-43		
Matches	446	Conservative	0	Mismatches 328	Indels 15
				Gaps	
QY 277	GCGCACTGGACCATGTCGCCATCGTCGGCGCGCGGCGCTCGTGCCCGGAATGCGGGG	336			
Db					
QY 3339	GTCGATACGGGATGCGCGCTATCGTCGGCGCGCGGATTTGTTGCCGACTTTGGCCCGG	3398			
Db					
QY 337	CTGAAACAGCGGCGCCAGCGTCGGCGCTCGCCAAAGAAAGCTCTCGTAGCGGACGGGCA	396			
Db					
QY 3399	GCCAAAGCGCGGCAAAACCGTGCTGTGGCCCAAGAGAAAGCTTGGTGAATGCGGACAA	3455			
Db					
QY 397	CTCTGATGCGGAGCGGCGCCAGAGAAACGCGCGCACATCCTCGCGGTGAGACGAGCAC	456			
Db					
QY 3459	ATCTCAATGACGAGCGCGTCACAGAAATTCGCGCGCTGTGTGTGTCGCGAATACAGACGACAC	3518			
Db					
QY 457	TCGCGGCTCTTTCAGGCGCTGGCGGGCGAGGACAGCGCTGC-----GTC 501				
Db					
QY 3519	AACGCCATCTTTCAGTGCATGACGGCGGGGTTAATGCCACAGGCCATCACCAAGAGGG	3578			
Db					
QY 502	GAGGCGCTACATCACAGGGGCTCGCGGGGCGCTTCGCGAGCTCGAGGCTCGAGGCGATC	561			
Db					
QY 3579	GCGCGCATTTTATGACCGCTCCCGGTGGCCCAATTCGACGGACGCGGATTAAGACGTTG	3638			
Db					
QY 562	CGCGCTCGACCGCTGGCGCGGAGGGGCGAGGCCATCCCACTGTGCATGCGGCGGAGCGATC	621			
Db					
QY 3639	TCGACGGCGACGCGCGGATCAGCGCGGTTGCCCATCTAAATGGGACATGGGGGCGAAGATT	3698			
Db					
QY 622	TCGATCGACAGCGCTCTGATGTTCAACAGAGCGCTGAGCTGATCGACAGCGCGCAATTC	681			
Db					
QY 3699	TCGGTGCATTTCGCGCACCATGATGAAGCAAAAGGCTCTGAAGTGTACAACTGATCGAACCTGCTTGTTG	3758			
Db					
QY 682	TTGGGCTTCGAGCGCGGACCGGATCGAGCGGCGTGTCCATCCGAATCCATGCTCCATGG	741			
Db					
QY 3759	TTCAACATGAGCCCGACCAAGATTGAAGTGTATTTCATTCGACAGCATCTTAATTCG	3818			
Db					
QY 742	ATGGTGGGCTTTCGACGGGGGCTGATGGCCCATCTCGGCGCGCGCGACATGCGCAC	801			
Db					
QY 3819	ATGGTGAATATGTCGATGGTTTGCGTTTGCGSCAGATGGTAATCCGACATGCGCACG	3878			
Db					
QY 802	GCCATCGGATTTCCGCGCAACTGCGCGGGGTCGGGGGAGAGTCCCGTCCCGGATCGAC	861			
Db					
QY 3879	CCGATACCGACGCGGATGGCTCTGGCGGGAAGCTTTTGACTGTGGTGGCGCGCTGGAAT	3938			
Db					

```

QY 862 CTCGCACGAGTTCGAGCCCTACCTTCAGAGCCTGAAGAGAGCTTTCCGGCCCTG 921
Db 3939 ATTTTCGAGAGGAGGCGACATGATTTGAAAAACCCGACTTGAACGGTTCTGCTG 3998
QY 922 AGGCTTGGCGAGAGCTATGATGCGCGCGCGCCGTGTGGGGCGCCGCTTCAACCGCCG 981
Db 3999 AGATTGGCTTATGAAAGCATCAAGCTGTGTGAATTATGCCAACGGTATGAAAGCAGCC 4058
QY 982 AAGGAGATCGCGCTGCATTCATTCGCGCGAGCAGCATGGGTTTCTGACATGCGCGCG 1041
Db 4059 AATGAATTTGCTGTCGAGCGCTTTTAATGAAGATCAATTCACATCGACATCGCGCTC 4118
QY 1042 GTGGTCGAG 1050
Db 4119 ATCATCGAG 4127

RESULT 5
US-09-712-363-112
: Sequence 112, Application US/09712363
: Patent No. US20020164588A1
: GENERAL INFORMATION:
: APPLICANT: Eisenberg, David
: APPLICANT: Rotstein, Sergio H.
: APPLICANT: Marcotte, Edward M.
: TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
: FILE REFERENCE: 07419-032001
: CURRENT FILING DATE: 2000-11-13
: PRIOR APPLICATION NUMBER: PCT/US00/02246
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: 60/179,531
: PRIOR FILING DATE: 2000-02-01
: PRIOR APPLICATION NUMBER: 60/117,844
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/118,206,
: PRIOR FILING DATE: 1999-02-01
: PRIOR APPLICATION NUMBER: 60/126,593
: PRIOR FILING DATE: 1999-03-26
: PRIOR APPLICATION NUMBER: 60/134,093
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 60/134,092
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 60/165,124
: PRIOR FILING DATE: 1999-11-12
: PRIOR APPLICATION NUMBER: 60/165,086
: NUMBER OF SEQ ID NOS: 292
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 112
: LENGTH: 1311
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
US-09-712-363-112

Query Match 18.2%; Score 215.6; DB 9; Length 1311;
Best Local Similarity 54.2%; Pred. No. 2.4e-41;
Matches 462; Conservative 0; Mismatches 384; Indels 6; Gaps 1;

```

```

Db 462 GGTCCCGGTGTTGCTGCTGCTGCGCGCGCGG-----CCGGTCAGATGTCCTCC 515
QY 441 GGTGAGACGCGAGCATCCGCGGCTTTTACGCGCTGCGCGAGCAGACAGCGCTGCTG 500
Db 516 GGTGACATCCGACACATCCGCGCTGCGCGAGCTGCTGCGCGCGAGCTTCCGAGAGCT 575
QY 501 CGAGCGCTCATCATCAGCGGCTCCGCGCGCGCTTCCGCGAGTGAAGCCTGAGCGCAT 560
Db 576 CGCCAGCTGTGTGAGCGCGCTCGGAGCGCGCTTTCGGGCGCTGTCGCGCGCGAGCT 635
QY 561 CCGCGCTGACCGCTGCGCGAGCGCGAGCCCATTCACATGCTGATGCTGCGCGAGCAT 620
Db 636 CGAGCATGTCAACCCCGAGCGAGCTGCGCGCATCTCACTGATGATGAGCGCGCATGAA 695
QY 621 CTCATCGACAGCGCTGATGATTCACAGAGCGCTCGAGCTGATCGAGACGCGGATTT 680
Db 696 CACGCTGAATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
QY 681 CTTCGCTTTCGAGCGCGAGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740
Db 756 GTTCGCTTTCGAGCGCGAGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
QY 741 GATGCTGCGCTTTCGAGCGAGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
Db 816 GATGCTGCTTTCATGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 875
QY 801 CGCATCGGATTCGCGCTGAGCTGCGCGCGCGCTGCGCGAGCTGCTGCGCGCGAGTGA 860
Db 876 ACCGATTTTCGTTAGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGGA 935
QY 861 CCTGCGACAGATTCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
Db 936 TTTCCATACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
QY 921 GAGCTTTCGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
Db 996 CGAGTTGCTGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
QY 981 CAAGAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
Db 1056 GAACGAGAGAGAGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1115
QY 1041 GGTGTCGAGA 1052
Db 1116 CATATCGCGA 1127

RESULT 6
US-09-974-300-1083
: Sequence 1083, Application US/09974300
: Patent No. US20020146721A1
: GENERAL INFORMATION:
: APPLICANT: Berka, Randy M.
: APPLICANT: Clausen, Ib Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: FILE REFERENCE: 10085.500-US
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: US/09/974,300
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 09/680,598
: PRIOR FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1083
: LENGTH: 1115
: TYPE: DNA
: ORGANISM: Bacillus licheniformis
US-09-974-300-1083

Query Match 15.2%; Score 179.2; DB 10; Length 1115;

```

Best Local Similarity 50.0%; Pred. No. 6,4e-33;
Matches 535; Conservative 0; Mismatches 523; Indels 12; Gaps 3;

QY 2 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 61
 DB 2 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 61
 QY 62 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 121
 DB 62 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 115
 QY 122 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 181
 DB 116 CCGAAAAACCGGAGAAATCATTTAAACCTTTAAACCGAAGATATGAGCGCGGGGATG 175
 QY 182 AGAGTGCCTGCGCGCGCTGCGGAGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAG 241
 DB 176 AGCATATCATATGAAACATTAAACAGCATTTCTTCTTCAACCTTTAAACAGGATCG 235
 QY 242 GGGCGAGGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 298
 DB 236 GGGAGAGAGCTTTATTCGAGCGGCGGCTTATCCCGAGCGGATATCGTCTCATGCTC 295
 QY 299 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 358
 DB 296 TCGGCGAGCTTCGATCTTTGGGGCCACCGGCTCCATCGCCGATCCACCTCGACCTCG 355
 QY 359 CCGTGGCCAAAGAAAGAACTTCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 418
 DB 356 CCGTGGCCAAATGAAAGAACTTCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 415
 QY 419 AGAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478
 DB 416 AATGAGATTCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 475
 QY 479 CCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 538
 DB 476 AGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 535
 QY 539 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 598
 DB 536 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 595
 QY 599 ACTGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
 DB 596 ATTGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 655
 QY 659 ACTGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 718
 DB 656 AGGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 715
 QY 719 ATCCGATTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
 DB 716 ACAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 775
 QY 779 TCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
 DB 776 TCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 835
 QY 839 AGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 895
 DB 836 CCGTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 895
 QY 896 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955
 DB 896 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955
 QY 956 TCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1015
 DB 956 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1015
 QY 1016 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065
 DB 1016 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065

DB 1016 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065

RESULT 7

US-09-738-626-2209
 ; Sequence 2209, Application US/09738626
 ; Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/37484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 2209
 LENGTH: 1176
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum

US-09-738-626-2209

Query Match 12.3%; Score 145.8; DB 9; Length 1176;
 Best Local Similarity 49.9%; Pred. No. 3,4e-25;

Matches 397; Conservative 0; Mismatches 392; Indels 6; Gaps 1;

QY 276 GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 335
 DB 285 GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 344
 QY 336 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
 DB 345 TCGGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404
 QY 396 ACTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455
 DB 405 GTTGTACCTCAAGAGAAAGCTG-----GGGCGAGATCATTCGGTGGATTCGAGGAG 458
 QY 456 CTCGCGGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 515
 DB 459 CTCGCGGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 518
 QY 516 CAGGCGGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 575
 DB 519 GAGAGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 578
 QY 576 GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 635
 DB 579 CCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 638
 QY 636 CTCGATGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 695
 DB 639 CAGCCTTATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 698
 QY 696 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 755
 DB 699 GAGTCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 758
 QY 756 CCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 815


```

RESULT 10
US-10-047-412A-28
Sequence 28, Application US/10047412A
Publication NO. US20020197656A1
GENERAL INFORMATION:
APPLICANT: Levin, Joshua Z.
APPLICANT: Budziszewski, Gregory J.
APPLICANT: Potter, Sharon L.
APPLICANT: Megrich, Lynette M.
TITLE OF INVENTION: Herpicide Target Genes and Methods
FILE REFERENCE: PB/5-30780DIV
CURRENT APPLICATION NUMBER: US/10/047,412A
CURRENT FILING DATE: 2002-04-11
NUMBER OF SEQ ID NOS: 29

```

OY	293	CGGCGATCGTGGGGGCGCGGGCGCTGTGGCCGGAATGGGGGCGGTGAAGCAGCGCCGCA	352
Db	533	CCGGAAATGTAAGTGTGTGGCCGGACTGMAGCTTACGGTTCTGCAATTTGACAGCAAGAAAG	592
OY	353	CGCTGGGCGCTCGCCACMACAAGAAAGCCCTCGTGCAGGACAGGAGCAACTCGTATCGGACGG	412
Db	593	ACATTTGCTTTGCAACAAGAGACATTAATGCAGAGTGTCCTTTGCTTCCGCTTGG	652
OY	413	CCGACAGAAACGGCGCCACATCTCTGCGGTGGACACGACAGCACTCCGGGCTTTCCAG	472
Db	653	CCAAACAACATTAATGATTAATAATCTTCGCGCAATTTCAACAAATCTCGCATATTTTCAGT	712
OY	473	CGGTGGGGGGGAGAGACACGGCCCGGTGAGCGGCGTCATCAACAGGGGTCGGGCGGAC	532
Db	713	GTAATTCAGATTTTTCCTGGAAGGCGCTCTGCGCAAGATAATTTGATGTCATCTTGATGAG	772
OY	533	CGTTCCGCGACTGAGACCTCGAGCGCATCCGCGCCCTGCACCGTGGCGGACGACAGGCC	592
Db	773	CTTTTAGGAGATTTGGCTGTGCAAAAGCTAAAGAAAGTTAAAGTAGGGATGGTTGAAGC	832
OY	593	ATCCCACTGCTCCATGGGCGCAGCGGATCTCCATTCGACAGCGCTTGATGTTCAACAAG	652
Db	833	ATCCAAACTGGAACTGAGGAAAGAAATACTCTGAGACTCTGTAAGCTTTTTCACCAAG	892
OY	653	CGCTCGAGCTATGAGAACGCGCGGAATCTTCGGCTTCGACGCGGACCGGATCGAGGGG	712
Db	893	GCTCTTGAGGTATTTGAAGGCCATTAATTTGTTTGAGACTGATATGACATATAGAGATTG	952
OY	713	TGCTGCAATCCGCAATTCATGCTGCATGCGATGGTGGGCTTCTGACAGGGGGCTGATGG	772
Db	953	TCAATTCATCCCAAGATATCATATCATATGATTAACACAGAGATTCATCTGTGCTTG	1012
OY	773	CCCATCTGGGCGCCCGGACATGGCCAGCGCATCGGATTCGGCGTGAACGTGGCGGCTC	832
Db	1013	CTCAATTTGGGTTGGCTCGAATATGCTGTTTACCGATTCCTTACACCATGTCATGAGCCGATA	1072
OY	833	GGGG-----CGAGGTCCCGGTGGCGCGGATGACATCCGCAAGATTTGGACAGCCCA	883
Db	1073	GAGTTCTCTTGTTCGAAGTAACTTGGCGAGACTTAACCTTTGCAAACTCGGTTATTGGA	1133
OY	884	CTTTCCAGAAAGCTGACAGAGAAACGCTTTCGCGCCCTGAGGGCTTGGCGAGACGTATGG	943
Db	1133	CTTTTCAGAAACCGAACAATGTGAATATCCCATCCATGGAATCTTGTATGCTGCTGAC	1192
OY	944	CGGCGCGCGCTGTGCGGGCGCGCTTCAACGCGGCGCAAGAGATCGGCTCGATCATTT	1003
Db	1193	GAGCTGAGGCGACATATGACTGAGGATCTTCAGCGCCGCAATGAGAAAGCTGTTGAAATGT	1252
OY	1004	TCATCGCGCGACGATCGGCTTCTTGACATGGCGGCGGTGTCGA	1049
Db	1253	TTATTTGAGAAAGATTAAGCTATTTTGGATATCTTCAAGGTGTGGGA	1298

US-09 923-556-1
; Sequence 1, Application US/09923556
; Patent No. US20020058321A1
; GENERAL INFORMATION:

```

; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Rieth, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
; FILE REFERENCE: 202005
; CURRENT APPLICATION NUMBER: US/09/923,556
; CURRENT FILING DATE: 2001-08-07
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/449,335
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: DNA encoding 1-deoxy-D-xylulose-5-phosphate
; OTHER INFORMATION: reductoisomerase
US-09-923-556-1

```

```

Query Match      10.5%; Score 123.6; DB 10; Length 1732;
Best Local Similarity 48.7%; Pred. No. 4.5e-20;
Matches 373; Conservative 0; Mismatches 384; Indels 9; Gaps 1;

```

```

QY 293 CGGCCATCGTGGGCGCGCGGCTCGTCCCGGATGCGGCGCTGAGCAGCGCCGCA 352
DB 550 CCGGAATAGTAGTGTGGCGGACTAAAGCTTACGTTGCTGCAATTAAGCAGGAAG 609
QY 353 CGGTGGCGCTCGCACAAGAAAGCCTCGTACGCGGCAAGCCTGATGCGAGCG 412
DB 610 ACATGCTCTTGCACAAGAAAGACATTAATCGAGGTGCTCTTGTGCTTCCGCTTG 669
QY 413 CCCAGGAAAGCGCGCAGCATCTCGGCTGAGACAGCAGCAGCTCGCGCTTTGAGG 472
DB 670 CCAACAACATTAATGTAAATCTTCCGCGACATTCAGAACTTCCATATTTCAGT 729
QY 473 CGGTGGGCGGAGGAGACGGCTGCTGAGCGGCTCATCATCAGCGCTCGCGCGGCG 532
DB 730 GATTCAGAGTTTGGCTGGAAGGCGCTGCGCAAGATTAATCTTGCATCTGCTGGAG 789
QY 533 CGTTCCGCGATGAGCGCTGAGGAGCATCCGCGCTGACCGTGGCGAGCGAGCGCC 592
DB 790 CTTTAAAGGATTTGGCTGTGAAAGAGTAAAGAAAGTTAAAGTAAAGCGATCGTTGAGC 849
QY 593 ATCCCAACTGCTCCATGCGGCGAGGATTCATCGACAGCGCGCTCGATTTCAACAGG 652
DB 850 ATCCCAACTGGAACATGGAAGAAATCACTGTGACCTCTGACCTTTTCAACAGG 909
QY 653 CGCTGAGCTGATGAGAGCGCGGAATTCCTGCGCTTGCAGCGCGAGCCGATGAGCGG 712
DB 910 GTCTTGGGTCATGGAAGCCATTAATTTGTTGAGCTGAGTATGAGATATAGAGATTG 969
QY 713 TCGTCCATCCGCAATCGATGCTCATGATGAGGCTCTGCGAGCGGGGCGCTGATG 772
DB 970 TCATTCATCGCAAGATATATACATTCATGATGAAACAGAGATTCATCTGTGCTTG 1029
QY 773 CCCATCTCGGCGCGCGGATGCGGCGCATCGGATTCGCGCTGAACTGCGCGGCTC 832
DB 1030 CTCATTTGGGTTGGCGCATATGCTTTACGATTTCTCTACACCATTCATGCGCCGATA 1089
QY 833 GCGGCG-----GAGGTCCCGCTGCGCGCGGATCGACCTCGACAGATTGCGAGCTCA 883
DB 1090 GACTTCCTGTTCTGAGATTAAGTGGCCAAAGCTTGAAGCTTTGCAAACTCGTTTCATTTGA 1149
QY 884 CTTTCCAGAGCCTGAGAGAGAGGCTTTCCGCGCTGAGAGCTTGGCGAGAGCTATGG 943
DB 1150 CTTTCAAGAAACGACGAAATGTGAATGCCATCCATGATCTTGTATGCTGCTGAGC 1209

```

```

QY 944 CGGCGCGGCGCTGTGCGGCGCGCGCTTCAACGCGCAAGAGATGCGCTGATCAT 1003
DB 1210 GAGCTGAGGCGACAAATGACTGAGATTTCTACAGCGCGCCCAATGAGAAAGCTGTGAATGT 1269
QY 1004 TCATTCGCGGAGCATCGGCTTCTGACATGAGGCGGCTGTGCA 1049
DB 1270 TCATTCATGAAAGATTAAGCTATTGGATATCTTCAAGGTTTGCA 1315

```

RESULT 12

```

US-10-047-412A-9
; Sequence 9, Application US/10047412A
; Publication No. US20020197696A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Weiglich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780DIV
; CURRENT APPLICATION NUMBER: US/10/047,412A
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
US-10-047-412A-9

```

```

Query Match      10.3%; Score 122; DB 9; Length 1434;
Best Local Similarity 48.6%; Pred. No. 1.1e-19;
Matches 372; Conservative 0; Mismatches 385; Indels 9; Gaps 1;

```

```

QY 293 CGGCCATCGTGGGCGCGCGGCTCGTCCCGGATGCGGCGCTGAGCAGCGCCGCA 352
DB 533 CCGGAATAGTAGTGTGGCGGACTAAAGCTTACGTTGCTGCAATTAAGCAGGAAG 592
QY 353 CGGTGGCGCTCGCACAAGAAAGCCTCGTACGCGGCAAGCCTGATGCGAGCG 412
DB 593 ACATTTGCTTTGCAAAAGAAAGACATTAATCGAGGTGCTCTTGTGCTTCCGCTTG 652
QY 413 CCCAGGAAAGCGCGCAGCATCTCGGCTGAGACAGCAGCAGCTCGCGCTTTGAGG 472
DB 653 CCAACAACATTAATGTAAAGATTTCTTCCGCGACATTCAGAACTTCCATATTTCAGT 712
QY 473 CGGTGGGCGGAGGAGACGGCTGCTGAGCGGCTCATCATCAGCGCTCGCGCGGCG 532
DB 713 GATTCAGAGTTTGGCTGGAAGGCGCTGCGCAAGATTAATCTTGCATCTGCTGGAG 789
QY 533 CGTTCCGCGATGAGCGCTGAGGAGCATCCGCGCTGACCGTGGCGAGCGAGCGCC 592
DB 773 CTTTAAAGGATTTGGCTGTGAAAGAGTAAAGAAAGTTAAAGTAAAGCGATTCGTTGAGC 832
QY 773 ATCCCACTGCTCCATGCGGCGAGGATTCATCGACAGCGCGCTCGATTTCAACAGG 652
DB 833 ATCCCACTGGAACATGGAAGAAATCACTGTGACCTCTGACCTTTTCAACAGG 892
QY 833 CGCTGAGCTGATGAGAGCGCGGAATTCCTGCGCTTGCAGCGCGAGCCGATGAGCGG 712
DB 893 GTCTTGGGTCATGGAAGCCATTAATTTGTTGAGCTGAGTATGAGATATAGAGATTG 952
QY 713 TCGTCCATCCGCAATCGATGCTCATGATGAGGCTCTGCGAGCGGGGCGCTGATG 772
DB 953 TCATTCATCTCAAGATATACATTCATGATGAAACAGAGATTCATCTGTGCTTG 1012
QY 773 CCCATCTCGGCGCGCGGATGCGGCGCATCGGATTCGCGCTGAACTGCGCGGCTC 832
DB 1013 CTCATTTGGGTTGGCGCATATGCTTTACGATTTCTCTACACCATTCATGCGCCGATA 1072

```


QY 515 TCAGGCGCTCCGGGCGGCGCTTCGCGACATGAGCGATCCCGCTGACCG 574
Db 701 GCGCCCTGGCGCGAGAGTGTTCGCGAGCGCGGCGGACCGCGCGCGCGCG 760
QY 575 TGGCCGAGGCGCGACGCGATCCCACTGCTGCAT 608
Db 761 AGCGGCGCGCGACCGCGCGCTGCGCGCTT 794

RESULT 15

US-09-860-846-32
; Sequence 32, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 11220
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-32

Query Match 5.6%; Score 66.6; DB 9; Length 11220;
Best Local Similarity 47.3%; Pred. No. 5.9e-07;
Matches 235; Conservative 0; Mismatches 259; Indels 3; Gaps 1;

QY 156 GCGGAGCTTCCGTCACCGCGCATGAGACTGCTGCCCGCTGCGGAGCGCTGCC 215
Db 1707 GGTGGCGTTCGTTCCTCCGTCAGGGCAGCAGTGGCGCGGATGGCGCGCACTCT 1766
QY 216 CGGAGCGGCGACCGAGGTCGGGCGGCGCGCGCATCGCGCGCGCGCGCGCGCC 275
Db 1767 CGAGGTGTCAGAGAGTTCGGGCGCGCGCATGCGCGAGAGCGCGCTCTCCGCTA 1826
QY 276 GCGCGACTGACATGTCGCGCATGTCGGCGCGCGCGCGCTGTCGCGGAGTGC 335
Db 1827 TGTGACTGTGTCGCGCGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1886
QY 336 GCTGAGCAGCG 395
Db 1887 GGTGAGCGTTCG 1946
QY 396 ACTCTGATGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 455
Db 1947 GCACGAGCG 2003
QY 456 CTCGCGGCTTTCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 515
Db 2004 CGCGTAGCTGCG 2063
QY 516 CAGCGCGTCCG 575
Db 2064 CAACTTCATGCG 2123
QY 576 GCGCGAGCG 635
Db 2124 GGAAGCCACCG 2183
QY 636 CTCGATGTTCAACAAG 652
Db 2184 CCCACCG 2200

Search completed: April 22, 2003, 06:16:53
Job time : 3994 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd..

OM nucleic - nucleic search, using sw model

Run on: April 22, 2003, 01:28:39 ; Search time 1961 Seconds
(without alignments)
9761.890 Million cell updates/sec

Title: US-09-673-198-31

Perfect score: 1182
Sequence: 1 atgcgcagcctgtcgcattt.....gtctccgcagcagcaaaag 1182

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	205	17.3	570	13 B1721015	B1721015 1031053EO
2	153.8	13.0	280	17 A2577432	A2577432 10803 Sho
3	129.4	10.9	687	10 BE213317	BE213317 EST0074 T
4	127.4	10.8	488	17 BH001147	BH001147 p12-F2 P1
5	124.4	10.5	626	13 BE519766	BE519766 HV-CE002
6	118.4	10.0	680	13 BU471224	BU471224 BU471224

C	7	113.8	9.6	784	13 B1751397
	8	112	9.5	1838	11 AY109535
	9	110.6	9.4	4646	12 BF428515
	10	109.4	9.3	543	13 BM406343
	11	109.4	9.3	594	14 B0518501
	12	107.4	9.1	687	10 AM775985
	13	105	8.9	838	12 BG588663
	14	98.8	8.4	798	10 BE412040
	15	97.2	8.2	712	12 BG586951
	16	93.8	7.9	619	13 B1470380
	17	91.6	7.7	469	13 B1894320
	18	91.6	7.7	546	13 B1674464
C	19	91	7.7	706	17 CNS010LS
	20	90.2	7.6	584	13 B1942645
	21	90.2	7.6	845	12 BF261910
	22	89.6	7.6	682	12 BG840487
	23	88.8	7.5	831	12 BG343005
	24	88.6	7.5	598	10 AV934537
	25	86.4	7.3	471	10 AM616251
	26	86	7.3	572	13 B1943128
	27	85.6	7.2	575	13 B1943466
	28	85	7.2	548	14 B0039624
	29	84.6	7.2	594	10 AV933746
	30	84.6	7.2	829	13 BM112204
	31	82.6	7.0	751	12 BF005132
	32	82	6.9	468	13 B1923028
	33	81	6.9	591	9 A1993895
	34	81	6.9	637	12 BG043747
	35	80.8	6.8	353	9 AL819484
	36	80.4	6.8	555	13 BM524138
C	37	80	6.8	925	17 CNS0091P
	38	79.8	6.8	675	12 BG449020
C	39	79.4	6.7	935	17 CNS006XK
	40	79.2	6.7	286	10 AM067754
	41	78.8	6.7	562	13 B1472011
	42	78.8	6.7	596	14 BQ295914
	43	78.8	6.7	670	12 BG453943
	44	78.8	6.7	692	12 BF598623
	45	78.8	6.7	743	14 BQ917023

ALIGNMENTS

B1751397 Ta01_20b0
AY109535 zea mays
BF428515 WHE1409_A
BM406343 EST580670
B0518501 EST625916
AM775985 EST335050
BG588663 EST490472
BE412040 J1L001.H0
BG586951 EST488720
B1470380 B1470380
B1894320 ga24e11.Y
B1674464 san8a05.
AL153680 Anopheles
B1942645 sk22d12.Y
BF261910 HV-CEA000
BG840487 MEST10-F0
BG343005 HVSMB9000
AV934537 AV934537
AM616251 EST307290
B1943128 sa43f10.Y
B1943466 sp73606.Y
B0039624 g601a03.Y
AV933746 AV933746
BM112204 EST559740
BF005132 EST433630
B1923028 EST542932
A1993895 701515340
BG043747 sv27f05.Y
AL819484 AL819484
BM524138 sal07f08.
AL053013 Drosophila
BG449020 NF005H11
AL066051 Drosophila
AM067754 SP0144 KR
B1472011 san98c10.
BQ295914 sa028f09.
BG453943 NF096C07L
BF598623 sv1907.Y
BQ917023 QHB20A15.

RESULT 1
LOCUS B1721015 570 bp mRNA linear EST 19-SEP-2001
DEFINITION 1031053EO4.y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B1721015
VERSION B1721015.1 GI:15696710
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 570)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.F., Strager,J., Sillow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES
source Location/Qualifiers
1..570
/organism="Chlamydomonas reinhardtii"

```

/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda zap II"
/notes="Vector: Bluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into Lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. Bluescript II SK- plasmids were excised from the
lambda zap clones by superinfection with ExSist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldi et al., (1996) Genome
Research 6: 791-806."

BASE COUNT      108 a      211 c      161 g      87 t      3 others
ORIGIN
Query Match      17.3%; Score 205; DB 13; Length 570;
Best Local Similarity 63.7%; Pred. No. 1.4e-27;
Matches 310; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 366 CAACAGAGAAAGCCTCTGACGGCAGGCAACTCTGATGCGAGCGCCAGAGAAAGG 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 CACAGAGGAGAACCTCATGTCGGGGGCGCCCTTCATCTCGCGGTGGCCAGAGTAGCG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 CGGCACATCTCGCGGTGACACGACGACACTCGCGGTCTTTAGGCGCTGGCGGGCA 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CTCGAATCTCTGCGCGCCGACAGCAGCAGCAGCGCATCTCTCAGGTGAGAGGGCCT 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 GGAACAGCGCTGCGTGCAGCGCGTCATCATCAGCGCGTCCGGCGGGCGTCCGCGACTG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GCCCGAGGGGGCGCTGGCGCGCATCTCTACCGCTCGGGGAGCGCTTCGCGACTG 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 GAGCCTGAGCGGATCGCGCGCTGACCGTGGCCGAGCGCAGCCATCCCAACTGCTC 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 GCCCGTGAAGAGTGGGAGAGTGAGCGGTGGCCAGCCAGCAGCCCAACTGCTC 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 606 CATGGCGCAGCGATCTCTCATCAGCAGCGCTCGATGTTCAACAAGCGCTCGAGCTGAT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 CATGGGCAAGATGACCGCTGACAGCAGCACTCATGAAGAGGCTCTGGAGGTGAT 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 666 CGAAGACGCGGAATTTCTGCGCTTCGAGCGGACCGAGTGGGCGGTCGTCATCCGCA 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 TGAGGCTCACTACTGTTCCGAGTGGACTAGACACATGACATGCTATCCACCCCA 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 726 ATCCATGTCATGCGATGCTGCGAGCGGAGCGGCGCTGATGGCCCATCTCGGCC 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 GTCCATCATCTACATCTCATGATCGAGCAGACTCTCTNNTGCTGGCGCAGCTGGGCTG 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 786 CGCGAGATCGCGCAGCGCATCGGATCGGCTGAACTGGCGGGTGGCGCGAGGTGCC 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GCCCGAGATCGCGCTGCGCATGATGACACATGCTCTGGCCGAGAGTGGCGGTCA 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 846 CGTCGCC 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GGAGGCC 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
A2577432/c 280 bp DNA linear GSS 08-DEC-2000
LOCUS 10a03 Shot-gun genomic library of Rhizobium strain ANU265 Rhizobium
DEFINITION sp. NGR234 genomic clone 10a03, DNA sequence.
ACCESSION A2577432
VERSION A2577432.1 GI:11603701
KEYWORDS GSS.
SOURCE Rhizobium sp. NGR234.

```

```

ORGANISM Rhizobium sp. NGR234
REFERENCE Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium.
AUTHORS 1 (bases 1 to 280)
TITLE Virey, V., Rosenthal, A., Broughton, W.J. and Perret, X.
JOURNAL Genetic snapshots of the Rhizobium species NGR234 genome
MEDLINE Genome Biol. 1 (6), RESEARCH0014 (2000)
2114532
COMMENT Contact: Virginie Virey
Laboratoire de Biologie Moléculaire des Plantes Supérieures
Université de Genève
1 Chemin de l'Imperatrice, Chambesey/Genève 1292, Switzerland
Tel: +44(0)1603450000
Fax: +44(0)1603450045
Email: virginie.virey@obs.jc.ac.uk
Class: shotgun.

FEATURES
source location/Qualifiers
1..280
/organism="Rhizobium sp. NGR234"
/strain="ANU265"
/db_xref="taxon:394"
/clone_lib="10a03"
/clone_lib="Shot-gun genomic library of Rhizobium strain
ANU265"
/notes="Vector: M13; derivative strain of NGR234 cured of
pNGR234a"

BASE COUNT      53 a      83 c      91 g      53 t
ORIGIN
Query Match      13.0%; Score 153.8; DB 17; Length 280;
Best Local Similarity 73.2%; Pred. No. 2.4e-18;
Matches 197; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 571 ACCGTGCGGAGGCGGCGCCATCCCACTGATGTCATGCGGCGGCGGATCTCCATGAC 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 ACGCCGACATATGAGCGCGCCGCGCATCGCACTGATGATGGGCTGAAATCTCGATCGAC 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 AGCGCTCGATGTTCAACAAGCGCTGAGCTGATCGACAGCGCGGAATCTTGGCTTC 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AGCGCTTGATGTTCAACAAGCGCTGAAATGATCGAGCGCGCGCATCTCTTTCGCGCTG 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 GAGCGGACCGGATGAGGCGGCGTGCATCCGCACTGATCGATCGATCGATCGATCGATCG 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 CGCCCGACCAAGATCGAGGTGATGCTCATCTCCGATCGGTCTGATGATGATGATG 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 TTCTGCGAGCGGCGCTGATGAGCGCCATCTCGCGCGCGCGCGCATCGCGCATCGCG 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 TACAGCGAGGATGCGTGTGGCGACCTCGCGTGGCGCGCGCGCATCGCGCATCGCG 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 811 TTGCGCTGAACTGCGCGGCTCGCGGCGA 839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 TATGCGCTCACTATCCGAGCGCTGCGA 3
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
BE213317/c 687 bp mRNA linear EST 30-JUN-2000
LOCUS BE213317
DEFINITION EST0074 Triticum aestivum Lambda zap Triticum aestivum cDNA clone
JAL1D.H11.T3.5' similar to putative 1-deoxy-D-xylulose-5-phosphate
reductoisomerase, mRNA sequence.
ACCESSION BE213317
VERSION BE213317.1 GI:8845070
KEYWORDS EST.
SOURCE Triticum aestivum
ORGANISM bread wheat.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
Triticeae; Triticum.
REFERENCE 1 (bases 1 to 687)
AUTHORS Anderson, J.M., Williams, C.E. and Goodwin, S.B.
TITLE Analysis of an EST database reveals a probable CF2 resistance gene
homolog in wheat
JOURNAL unpublished (2000)

```


FEATURES		Location/Qualifiers
source		1. 680
		/organism="Hordeum vulgare subsp. vulgare"
		/cultivar="Haruna Nijo"
		/db_xref="taxon:112509"
		/clone="baall1f12"
		/clone.lib="K. Sato unpublished cDNA library, cv. Haruna Nijo adult, heading stage top three leaves"
		/dev_stage="adult, heading stage"
BASE COUNT		180 a 145 c 181 g 174 t
ORIGIN		
Query Match		10.0%: Score 118.4; DB 13; Length 680;
Best Local Similarity		49.9%: Pred. No. 5.8e-12;
Matches 333;		Conservative 0; Mismatches 326; Indels 9; Gaps 1;
QY	79	CCCCGAGCGTTCCGACCCGTCGCTTGACACGGCGGGCGCAACATCCGGCGATGCGCGAA 138
Db	9	CCGTATGAATGTCGCGGGTGTGTGCTCTACTGCTGCTCCAAATGTCACCTTCTACTGAT 68
QY	139	ATGGGCGCTGCGCTGAAGCGCGAGCTTCGCGACCGCGGACGGAGAGACTGCGTGGCCGG 198
Db	69	CAGGTGAAGACGTTCAACCAAGCTGGTTGCGCTTAAGAGAGCAATCTATTCTTAACAG 128
QY	199	CTGCGCGAGGGCGCTGCGCGG-----GACGGGACCGAGGTGCGGGGCGGGCGCGAG 249
Db	129	CTAAGGACGGTGTAGTGTGTGTGAGGAGATCGCGAAATATTCTTGCGAGCAAGT 188
QY	250	GCCATCGCCGAGGCGCGCGACCGCGCCGACCTGGACATGTGCGGCATGTGGGCGCC 309
Db	189	GTCATGAGGTTGCTGCGCGACCCAGATSCAGTTACAGTGTGTCGCGCATTAAGGTGT 248
QY	310	GCGGGCTCTGTGCCCGGAGATGCGGGCGGTGAAGACAGGCGCGACCGCTGCGTCCCAAC 369
Db	249	GCAGGACTCAGGCTCAGACAGTTGACAGTGTGAAGCTGGGAAGACATAGCACTGGCGAAC 308
QY	370	AAGGAAAGCCTGTGTACGCGCAGGCGCACTCTGTATGCGGACGCGCCAGAGAAAGCGGCC 429
Db	309	AAGGAGACACTTATTGTGACGGGCGCCGTTGCGTGTCCCTTGCGCACAGCAATGTG 368
QY	430	ACGATCTCGCGGTTGGACAGGAGCACTCCGGGGTCTTTACAGGCGCTGGCGGCGCAGGAC 489
Db	369	AAATATCTTCTGCTGATTCAGAGCACTCTGAAATATTTCASTGTATACAGAGCTTGTCT 428
QY	490	ACGGCCTCGTGCAGCGCGGTCAATCAACGGCGTCCGGCGGGCGCTTCGCGACGTGAGAC 549
Db	429	GAAGATCACTTCGTCGTGTTATCTTGACTGCGTGTGGGGGTCTTTCAGGACATGCGCA 488
QY	550	CTCGGCGCATCCGGCGCTGCACCGGTGGCGGGGCGGAGCGCCATCCCACTGTGTCCAG 609
Db	489	GTAGGAAGCTGAAGAGAGCTAAGAGTTGCGGAGTCTTTGAAGACCCCAATTGTAGAGCATG 548
QY	610	GCGCAGCGGATCTCCATCGACAGACCGCTCGATGTTTCAACAGAGCGCTCGAGCTGATCGAG 669
Db	549	GGGAAGAAATACACAGTAGATTCTGCTACTTATTATTCACAAGGGGTAGAAATTATCGAG 608
QY	670	ACGCGCGAATTTCTTGCGCTTCGAGCGCGGACCGGAGTGGCGGTGCTCATCCGAATCC 729
Db	609	GCACATTTTGTGTGTGCTGTAATATGATGACATTGACATTGTGATTCACCCGCAATCC 668
QY	730	ATCGTCCA 737
Db	669	ATCATACA 676
RESULT 7		
BI751397/c		BI751397
LOCUS		784 bp mRNA linear EST 25-SEP-2001
DEFINITION		Ta01_20b04_R
Tat01_AAFc_EORC_Fusarium_graminearum_inoculated_wheat_heads		
Triticum aestivum cDNA clone Ta01_20b04, mRNA sequence.		
ACCESSION		BI751397

VERSION	KEYWORDS	EST.
BI751397.1	GI:15773199	
SOURCE	ORGANISM	REFERENCE
bread wheat.	Triticum aestivum	1 (bases 1 to 784)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Spemolophyta; Triticum.	Ouellet, T., Dan, H., Koul, A., Chapados, J., Couroux, P., De Moors, A., Harris, L. J., Hattori, D. I., Robert, L. S., Singh, J. A., Sprott, D. and Tinker, N. A.
JOURNAL	COMMENT	FEATURES
Expressed Sequence Tags from Wheat Heads 24 Hours after Spray Inoculation with Fusarium graminearum	Unpublished (2001)	Location/Qualifiers
Contact: Ouellet, Therese	Eastern Cereal and Oilseed Research Centre	1..784
Agriculture and Agri-Food Canada	Needy Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA	/organism="Triticum aestivum"
Tel: (613) 759-1658	Fax: (613) 759-1701	/cultivar="Frontana"
Email: ouellet@erem.agr.ca.		/db_xref="taxon:4565"
		/clone="Ta01.20b04"
		/clone_lib="Ta01.ARC_ECORC_Fusarium_graminearum_inoculate"
		d_wheat_heads"
		/tissue_type="heads"
		/dev_stage="anthesis"
		/note="Vector: pCEM-T easy; site_1: EcoRI; site_2: EcoRI; Controlled chamber-grown wheat heads were spray inoculated at mid-anthesis with a Fusarium graminearum microconidial suspension (50,000 spores/ml) and kept under intermittent misting for 24 hours, then collected and immediately frozen in liquid nitrogen."
BASE COUNT	ORIGIN	
172 a	237 c	185 t 3 others
Query Match	9.6%	Score 113.8; DB 13; Length 784;
Best Local Similarity	53.2%	Pred. No. 3.9e-11;
Matches 292; Conservative	0; Mismatches 247; Indels 10; Gaps 2;	
522 GTCCGCGCGCCGTTTCCGCACTGAGCCCTCGAGCGCATCCGCGCCTCACCCTGACCGCA	581	
784 GTCTGGCGCGTCTTTCAGAGGACTACGAGTGAAGATGTAAGGTTGCCGA	725	
582 GCGCGAGCCCATCCCACTGTCATGCGCGGAGGATCTTCATCCGACGACGCTCGAT	641	
724 TGCCTTGAAGACCCAACTGAGCATGGGGAAGAAATACAGTAGATTCTGCTACTTT	665	
642 GTTC-AACAGGCGCTCAGCTGATCGAGCGCGGGAATTTTGGCTTCGAGCGGAC	700	
664 GTTCAACAGAGGGTTAGAGTTATCGAGGGCGCATTTATTGTTGGTCTGTAATGATG	605	
701 GGATCGAGGCGGTGTCATCCGCAATCTGTCATGCGATGATGGGCTTTCGACG	760	
604 ACATTGAGATTGTGATTCACCCACAGTCCATCATACACTATGATTTAAACCCAGGATT	545	
761 GGGGCTGATGAGCCCATTCGCGCCCGCGCGCATGCGCGACGCGCGGATTCGGCGGA	820	
544 CATCTGCTCTGAGCTCAGCTGGGATGCGGACAGACATGCGCTACCGCATCTTATACCTTGT	485	
821 ACTGCGCGGCTC-----GCGCGAGATGCGCCGTCGCGCGGATGCGACTTCGACAGA	871	
484 CTGCGCGAGCGAGGTACTGCTCCGAGGTCACCTGCGCCCGCGCTTACACTTTGCAAGC	425	
872 TTGCGAGCCTTACCTTCAGAGAGCTGACGAGAGAACGCTTTTCGGGCTGAGGCTTGGCG	931	
424 TGGGTGCTGACATTTAAAGCTCCCGCAACGATGAAATATACCATTCGAGGTGATCTCGCGT	365	

QY 932 GAGACGTCATGCGCGCGCGCTGTGCGGCGCCCTTCACGCGGCCAGAGATGC 991
 Db 364 ATGCGGCGAGGCGCGCGCTGACACCATGAGCGAGTTTGACGCGCTCAAGAGAG 305
 QY 992 CGCTCGATCATTTATCCCGGAGCGCATCGGCTTCTGACATGCGCGCGTGTGAGG 1051
 Db 304 CGGTGAGACTTTCATCGACGAAAGATGACCTACCTGACATCTTCAAGTGTGTGAGA 245
 QY 1052 AGACGCTCG 1060
 Db 244 TGACGTGCG 236

RESULT 8
 LOCUS AY109535 1838 bp mRNA Linear HTC 25-MAY-2002
 DEFINITION zea mays CL389_1 mRNA sequence.
 ACCESSION AY109535
 VERSION AY109535.1 GI:21213292
 KEYWORDS HTC.
 SOURCE zea mays.
 ORGANISM zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 Clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1838)
 Hainey,C.F., DoJan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1838)
 Coe,E.C.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA

FEATURES
 source
 Location/Qualifiers
 1..1838
 /organism="Zea mays"
 /db_xref="MaizeDB:631829"
 /db_xref="taxon:4577"
 /clone="CL389.1"
 /clone_lib="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACS in conjunction with the Maize
 Mapping Project"

BASE COUNT 438 a 398 c 433 g 438 t 131 others
 ORIGIN

Query Match 9.5% Score 112; DB 11; Length 1838;
 Best Local Similarity 44.9%; Pred. No. 8.3e-11;
 Matches 347; Conservative 0; Mismatches 416; Indels 9; Gaps 1;

QY 293 CGGCATCTGTGGCGCGCGCTGTGCGGAGATGCGGCGCTGAAGCAGCGCCCA 352
 Db 707 CAGGATAGTAGTTGTGCGAGGCTGAAGCTTCACTGCAATTAAGCTGTAAAG 766
 QY 353 CGCTGGCGCTGCGCAACAGAAAGCTCTGACGCGCAGGCAACTCTGTATGGGACGG 412
 Db 767 ACATAGATTGGCAACAAAGAGACACTTATTGACAGGTGCTTTGTCTCCCTTG 826
 QY 413 CCCAGGAGAGCGGCGCATCTGCGCGGAGACAGCAGCCTCCGCGTCTTTCAGG 472
 Db 827 CACCAACACCAAGTGAATTTCTTCAGCTGATTCTGACACCTCTGNNNNNNNNNN 886
 QY 473 CGCTGGCGGCGAGACAGCGCTGCGTGAAGCGCTCATCATCAGCGCGTCCGCGGCG 532
 Db 887 NNN 946

QY 533 CGTTCGAGCTGAGGAGCTTCGAGCGCATCCGCGCTGACCGGTGCGGAGCGAGCGCC 592
 Db 947 CTTTCAGGAGCTGGCGCACTTTCAGAGGCTGAAGAATGTAAAGTTGCTGACGCTTAAAGC 1006
 QY 593 ATCCCACTGCTCCATGAGGCGGAGATCTCATCGACAGCGCTTCATGTTCAACAAG 652
 Db 1007 ATCCAACTGGAATATGGGAAGGAAGATCACAGTATGCTTCACTTATTAACAAGG 1066
 QY 653 CGCTGAGCTGATCGAGAGCGCGCAATCTGCGCTTGACCGGAGCGGATCGAGCGG 712
 Db 1067 GTTTAGAGTTATTGAGACACATATTATTATTGCTGGAATATGATGACATTGATG 1126
 QY 713 TCGTCATCCGCAATTCATGCTCATGCGATGAGGCGCTTCTGCGAGCGGCGCTGATG 772
 Db 1127 TGATTACCCACAGCTCTCATACCTGATGTTGAACCCAGAGATTATCTGCTAG 1186
 QY 773 CCCATCTGCGCGCGCGCGCATGCGCCAGCATGCGATGATTTGGCGCTGACCTGCGCGG 829
 Db 1187 CTGATTTGGATGAGCGCAATATGCGGTTACCAATCTTATACCTTATCATGAGCGCAT 1246
 QY 830 -----GTGCGGCGAGGTGCGCGCGCGGATCGACCTCGACAGATTGCGAGCGCTCA 883
 Db 1247 GAATCTATTGCTCTGAGGTCTACCTGCGCGCTGATCTTTCGAGTTGGTTCACTGA 1306
 QY 884 CTTTCAGAACCTTGACAGGAGAGCGCTTTCGCGCGCTGAGGCTTTCGCGAGCGTCATG 943
 Db 1307 CATTGAGAGCTTCAGACAGCAACCTAAATACCATCAATGACCTGATGACCTGAGCC 1366
 QY 944 CGCGCGCGCGCTGTGCGGCGCGCGCTTCACAGCGCGCGCAAGAGATGCGGCTCATCAT 1003
 Db 1367 GCGCTNNNNNCACATGACAGAGTCTGAGCGCGCTGAAAGAGCGCGCTGAGTTGT 1426
 QY 1004 TCATGCGCGAGCAGCATCGGCTTTCGACATGCGGCGCGCTGAGAGAGAC 1055
 Db 1427 TCATTGAGAGAAATATGCTACCTGACATATTCAGATGAGTGTGAGCTTAC 1478

RESULT 9
 LOCUS BF428515 446 bp mRNA Linear EST 29-NOV-2000
 DEFINITION WHE1409_A03_A05S wheat drought stressed leaf cDNA library Trilicium
 aestivum cDNA clone WHE1409_A03_A05, mRNA sequence.
 ACCESSION BF428515
 VERSION BF428515.1 GI:11440279
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Trilicium aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 1 (bases 1 to 446)
 Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Han,P.S., Hsiao,C.C.,
 Kang,Y., Iazo,G.R., Miller,R., Nguyen H.T., Rausch,C.J., Seaton
 C.L., Tong,J.C. and Zhang,D.
 The structure and function of the expressed portion of the wheat
 genomes - Drought stressed leaf cDNA library
 Unpublished (2000)
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@wr.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.
 Location/Qualifiers
 1..446
 /organism="Trilicium aestivum"
 /cultiivar="TAM W101"
 /db_xref="taxon:4565"
 /clone="WHE1409_A03_A05"

```

/clone_lib="Wheat drought stressed leaf cDNA library"
/tissue_type="leaf"
/dev_stage="Full tillering stage"
/lab_host="E. coli S0LR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI. Plants were given a gradual
stress down to 65% and 78% RWC at Texas Tech University
(D. Zhang in Ht Nguyen lab). Total RNA and poly(A) RNA
were prepared, a cDNA library was made, and the cDNA
clones were in vivo excised to give Bluescript phagemids
in the T7 Close lab (Choi, Close) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."

BASE COUNT      112 a      103 c      108 g      123 t
ORIGIN

Query Match      9.4%; Score 110.6; DB:12; Length 446;
Best Local Similarity 53.9%; Pred. No. 1.5e-10;
Matches 227; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 419 AGACGCGCCGACATCTGCGGTGAGACGAGCAGCCTCCGGCTTTCAAGCGCTGG 478
    || || || || || || || || || || || || || || || || || || ||
DB 19 AGCACAATGTGAATAATCTTCTGCTGATTCAGAGCAGCTGTGCAATATTTCAGTATAC 78

QY 479 CGGCGAGGACACAGCGCTCGTGAGCGGCTCATCATCACGGGCGCGCGCTTCC 538
    || || || || || || || || || || || || || || || || || || ||
DB 79 AAGGCTTGTCTGAAGATCGCTCCGTCGGCTTATCTGACTGCTGCGGCGGTCTTCA 138

QY 539 GCGACTGAGAGCTCGAGCGCATCCGGCTCGACCGCTGGCGGAGGCGCCATCCA 598
    || || || || || || || || || || || || || || || || || || ||
DB 139 GGGACTGGCGAGTAGAAGCTGAAGATTAAGGTTGCGGATTCGTAAGACCCCAA 198

QY 599 ACTGTCATGGGCGCGCATCTCCATGCAGCGCTCGATGTTCAACAAGGCGCTCG 658
    || || || || || || || || || || || || || || || || || || ||
DB 199 ACTGAGCATGGGGAAGAAATCAGATGATTTCTCTACTTTGTTCAACAAGGCTTAC 258

QY 659 AGCTGATCGAAGCGCGAATTTCTTGCGCTTGAGCGCGGACCGGATGAGCGGCTGCC 718
    || || || || || || || || || || || || || || || || || || ||
DB 259 AAGTATTCGAGCGACATATTTGTTGTTGTCGTAATATGATGACATTTGATGATTTC 318

QY 719 ATCCGCAATCCATCTCCATGCGATGGGCTTTCGCAGCGGCGCTGATGCCATTC 778
    || || || || || || || || || || || || || || || || || || ||
DB 319 ACCCAGATGTCATCATCTCTATGATTGAACCCAGGATTCATCTGTCGTGGCCACG 378

QY 779 TCGGCGCCCGCAGATCGCCACGCGCATGTCGCGTGAACCGCGGCGCGGCGG 838
    || || || || || || || || || || || || || || || || || || ||
DB 379 TGGGATGGCTGACATGCGGCTACCAATCTTATACCTTTGCTTGCCAGATAGACTCT 438

QY 839 A 839
DB 439 A 439

```

RESULT 10
BM406343 543 bp mRNA linear EST 22-JAN-2002
LOCUS BM406343
DEFINITION EST580670 potato roots Solanum tuberosum cDNA clone cPR02706 5' end
/ mRNA sequence.

ACCESSION BM406343.1 GI:18257973
VERSION BM406343.1
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE
AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S.,
Ullrich, T., Chiemingo, A., Bougri, O., Buell, C.R., Romling, C.,
Tanksley, S. and Baker, B.
TITLE Generation of ESTs from potato roots
JOURNAL Unpublished (2001)

```

COMMENT
Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1..543
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPR02706"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="S0LR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley Lab.
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

BASE COUNT      148 a      119 c      121 g      155 t
ORIGIN

Query Match      9.3%; Score 109.4; DB:13; Length 543;
Best Local Similarity 54.0%; Pred. No. 2.4e-10;
Matches 224; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 415 CAGGACAGCGCGCCACGATCTCGCGGTGAGACGAGCAGTCCCGGCTTTACAGGG 474
    || || || || || || || || || || || || || || || || || || ||
DB 2 CACACACATTAAGTGAATTTCTCTCGCATTCAGAACATTCACCTATATTCAGTGC 61

QY 475 CTGGCGGCGAGGACACAGCGCTCGTGAGCGGCTCATCATCACGGGCGCGCGGCG 534
    || || || || || || || || || || || || || || || || || || ||
DB 475 CACACACATTAAGTGAATTTCTCTCGCATTCAGAACATTCACCTATATTCAGTGC 61

QY 535 TTCCGCGACTGAGCGCTCGCATCCGCGCTGACCGCTGCGGCGGCGGCGGCGCAT 594
    || || || || || || || || || || || || || || || || || || ||
DB 122 TTCCGAGGACTGGCGCACTTGGAAGTTGAAGAAGTTAAAGTACGATGCTTTGAAGCAT 181

QY 595 CCCAATCTGTCATGGGCGCAGCGGATCTCCATGCAGCGGCTCGATGTTCAACAAGCG 654
    || || || || || || || || || || || || || || || || || || ||
DB 182 CCCAATTTGAGAACATGGGAAAAAAGATTACGTTGATTCGCCCATTAATTAAGAGGCT 241

QY 655 CTGACCTGATGAGAGCGGCGCAATCTCGGCTTGAGCGCGGAGCGGATTCGAGCGGTC 714
    || || || || || || || || || || || || || || || || || || ||
DB 242 CTGGAAGTTATGAGAGCTACTACTCTTTGCGGCGCGAGATGACAAATGAAATGTC 301

QY 715 GTCCATCCGCAATCCATCTCCATGCGATGATGGGCTTCTGCGACGCGGCGCTGATGCC 774
    || || || || || || || || || || || || || || || || || || ||
DB 302 ATCCATCCCAATCCATCATATCATTCGATGCTGGAACACAGAGATTCATCAGTATGGCA 361

QY 775 CATCTCGCGCGCGCGCAGATGCGCCAGCGCATCGGATTCGCGCTGATGCGCGG 829
    || || || || || || || || || || || || || || || || || || ||
DB 362 CAACATAGGCTGCTGATATGCTTTGGCCATCCCTTAATACCTTAATGCGCGCG 416

```

RESULT 11
B0518501 594 bp mRNA linear EST 10-JUN-2002
LOCUS B0518501
DEFINITION EST625916 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMJMS4
5' end. mRNA sequence.

ACCESSION B0518501.1 GI:21377370
VERSION B0518501.1
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE
AUTHORS
TITLE
JOURNAL

AUTHORS Buelli, C. R., Hatt, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Rastepo, S., Giffitts, H., van der Hoeven, R., Tsai, J. and Karamycheva, S. A.

TITLE Generation of a set of potato cDNA clones for microarray analyses

JOURNAL Unpublished (2002)

COMMENT Other_ESTS: EST625517

QY 533 CGTTCCGCGACTGAGCCCTGACGCCGCTGACCCGTGCGGCGAGCGCCAGCC 592
 DB 419 CTTTGGGATTTGGCTGTGAGAACTGAAGATGTTAAAGTTCGATTAAC 478
 QY 593 ATCCCACTGTCATGCGCCAGCGATCTCCATCGACAGCGCTGATGTTCAACAAG 652
 DB 479 ATCCCACTGTCATGCGCCAGCGATCTCCATCGACAGCGCTGATGTTCAACAAG 538
 QY 653 CGGTGAGTGTGACGAGCGCGGATTCGTGGCTTCAGCGCGCGGATTCGAGCGG 712
 DB 539 GTCTAGAGTGTGACGAGCGGATTCGTGGCTTCAGCGCGCGGATTCGAGCGG 598
 QY 713 TGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 772
 DB 599 TGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 658
 QY 773 CCCATCTCGCGCCCGCGGATTCG 797
 DB 659 CACAAATGCGGCTGCGCTGATTCG 683

RESULT 13

LOCUS BG588663 838 bp mRNA linear EST 12-APR-2001
 DEFINITION EST90472 MHRP- Medicago truncatula cDNA clone PMHRP-57K17, mRNA
 sequence.

ACCESSION BG588663
 VERSION BG588663.1 GI:13606803

KEYWORDS

SOURCE

ORGANISM Medicago truncatula
 barrel medic.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Liliolidae;
 Medicago.

REFERENCE 1 (bases 1 to 838)
 Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
 and Fraser,C.M.

TITLE ESTs from phosphate-starved roots of Medicago truncatula, 2001
 JOURNAL Unpublished (2001)

COMMENT Contact: Harrison M.J.
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401
 Tel: 580-223-5810
 Fax: 580-221-7380

Email: mjharrison@noble.org
 The Samuel Roberts Noble Foundation: N387414e TIGR sequence name:
 MTHBC69YK More information is available at: <http://www.medicago.org>
 Seq primer: SKmod (CTA GAA GTA gta gat CC).

FEATURES
 Location/Qualifiers
 1..838

/organism="Medicago truncatula"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="PMHRP-57K17"
 /clone_1lb="MHRP-"
 /tissue_type="roots"
 /dev_stage="phosphate-starved"
 /lab_host="XLOLR"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; At the trifoliolate stage, M. truncatula plants were
 transplanted to phosphate-free sand and grown for a
 further 30 days. During this period, they were fertilized
 twice weekly with 1/2 Hoaglands solutions containing 20mM
 potassium phosphate. cDNA was prepared from polyA+
 enriched RNA. The cDNA was directionally ligated into the
 UniZap XR vector from Stratagene and packaged using
 GigaPack III Gold packaging extracts. Plasmids containing
 cDNA inserts were excised from the recombinant lambda-zap
 phage using Ex-assist helper phage and propagated in
 XLOLR cells."

BASE COUNT 225 a 163 c 186 g 264 t
 ORIGIN

Query Match 8.9%; Score 105; DB 12; Length 838;
 Best Local Similarity 48.4%; Pred. No. 1.5e-09;
 Matches 328; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 388 GCAGGCGCACTCTGATGCGAGAGCGCCAGAGAACGCGCCACGATTCCTGCGGTGAC 447
 DB 1 GAGAGTCCTTTGCTCTCTCTCTGCGCGGAGCAATATGTAATAATCTCTCTGAT 60
 QY 448 AGCGACACTCGCGGCTCTTTAGCGCGCTGCGCGCGAGAGACAGCGCTGCGTACGCC 507
 DB 61 TCCGAACATTCGCTCTTTTTCAGTGTATTCACAGGCGGTAAACAGAGGTCACATAGCGCA 120
 QY 508 GTCATATACAGCGCGCGCGCGCGCTTCGCGAGATGAGCTCGAGCGCATCCCGCGC 567
 DB 121 GTCATTTTAACTGATCCGAGAGGCTTTTAGGAGTGGCTGTGAGAACTGAAAGAT 180
 QY 568 TGCACCGTGGCGAGCGAGCGCGCCATCCCACTGCTCATGGCGCAGCGATCTTCATC 627
 DB 181 GTTAAAGTTGCGCATGCTATTAACATCCCACTGAGATGCGGAAAGATTAACGCTG 240
 QY 628 GACAGCGCTCGATGTTCAACAAGCGCTCGAGCTGATCGAGAGCGCGCAATTCCTGCG 687
 DB 241 GACTCTGTAACTTTTAAAGGCTTGAAGATGATGAAGCATTAATTAATTTGGA 300
 QY 688 TTCGACCGCGAGCGGATCGAGCGGCTGCTCCATCCCATTCATCTCATCGATGCTG 747
 DB 301 GCTGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 748 GCGCTTGGCGAGCGGCGCTGATGCGGCGCCATTCGCGCGCGCGCGCGCGCGCGCATC 807
 DB 361 GAACACAGATTCATCTGTTCTGACATAGGCGGCTGATGCGGTTTGCATTC 420
 QY 808 GGAATGCGCGCTGAACTGCGCGG-----TCGCGGAGAGTGCCTGCGCGCGATC 858
 DB 421 CTTTATACATTAATCAATGCGCTGACAGATTTATTTGTTGAAGTCACTTGGCTCGGCTT 480
 QY 859 GACCTCGCACATGTCGAGCGCTGACCGCTGACAGACCGTCAGAGAGAGCGTTTCGCGGC 918
 DB 481 GATCTTGCAGAGCTGTTGCTTCTTAAAGCTCGACAACTTAAGTATTCATCC 540
 QY 919 CTGAGGCTTGGCGGAGCGATGATGCGCGCGCGCTGCGCGCGCGCTTCAACGCG 978
 DB 541 ATGATCTTGGCTATGCTGCTGCGCGCGCGCGCGCGCGCATGACAGTGTCTTATGGA 600
 QY 979 GCCAGAGATGCGCTGATTCATTTTCATTCGCGCGAGCATGCGGTTTCTGACATGCG 1038
 DB 601 GCMAATGAGAGAGCTGTAGAGATGTTATGATGAAGATTAGCTATTTGATATTTTC 660
 QY 1039 GCGGTGTCGAGGAGAC 1055
 DB 661 AAGGTTGAGGCTAAC 677

RESULT 14

LOCUS BE412040 798 bp mRNA linear EST 24-JUL-2000
 DEFINITION JUI001.H07R990511 TREC JUL wheat leaf library Triticum aestivum
 cDNA clone JUI001.H07, mRNA sequence.

ACCESSION BE412040
 VERSION BE412040.1 GI:9409886

KEYWORDS EST.
 SOURCE bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 798)
 Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
 S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
 Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,D., Joudrier,P.,

Db 421 CTGGCCCTGGCTTGATCTTTGCAGCTGGTCCCTTACTTTAAAGCTCTGACACA 480
QY 905 AACGTTTCGGCCCTGAGGCTTGGCGAGAGCTCATGGCGGCGGCGCTGTGGGGC 964
Db 481 TTAGTATCCATCATGAATCTTGCTATGCTGTGGCGGCTGAGGCGCACCATTGACAG 540
QY 965 CCGCTTTCACGGCGGCAAGAGATGCGCTCGATCATTTTCATCGCCGAGCGCATCGGGT 1024
Db 541 GTGTCTTAGTGCAGCAATGAGAAAGCTGTAGAGATGTTTATGATGAAGAAAGATTAGCT 600
QY 1025 TTCTGACAT 1034
Db 601 ATTGGATAT 610

Search completed: April 22, 2003, 03:27:01
Job time : 1985 secs

PR 15-FEB-1999; 99JP-00357355

Methylinomona 16a O
Plasmid PRU575.dxs
Neisseria mening
Mycobacterium tube
Mycobacterium tube
Neisseria meningit
Neisseria gonorrhoe
Isoprenoid related
1-deoxy-D-xyliose
Propionibacterium
Isoprenoid related
Bacillus lichenifo
Isoprenoid related
C glutamicum codin
C glutamicum codin
C glutamicum codin
Polynucleotide seq
1-deoxy-D-xyliose
Listeria monococ
Listeria innocua c
A. thaliana 1-deox
A. thaliana DXPI
A. thaliana DXPI
Isoprenoid related
1-deoxy-D-xyliose
Corynebacterium gl
Arabisopsis herbic
1-deoxy-D-xyliose
Arabisopsis herbic
1-deoxy-D-xyliose
Isoprenoid related
DNA encoding novel
Chlamydia pneumoni
Nucleotide sequenc
Arabisopsis thalia

XX PA (KYOW) KYOMA HAKKO KOGYO KK.
 XX PA
 XX PI Miyake K, Hashimoto S, Motoyama H, Ozaki A, Seto H, Kuzuyama T,
 XX PI Takahashi S;
 XX DR WPI: 1999-620434/53.
 XX P-PSDB: AAY52839.
 XX
 XX Preparation of recombinant isoprenoid compounds useful for treatment of
 XX heart diseases, osteoporosis and hemostasis, preventing cancer and
 XX immunopotentialation
 PS Claim 7; Page 133-136; 145pp; Japanese.
 XX
 XX The present invention describes the preparation of an isoprenoid
 XX compound comprising using at least 1 DNA e.g. encoding proteins which
 XX elevate the efficiency of the synthesis or DNA encoding a farnesyl
 XX pyrolic producing enzyme. The method of preparation of an isoprenoid
 XX compound comprises using at least 1 DNA, a vector, cloned cells, their
 XX derived recombinant DNAs or transformed products in a culture system and
 XX extracting the isoprenoid accumulated in the medium. The DNA encodes at
 XX least 1 of the following: (a) a compound for activating or catalysing
 XX the production of 1-deoxy-D-xylulose-5-phosphate from pyruvic acid and
 XX glyceraldehyde-3-phosphate; (b) an enzyme producing farnesyl pyrolic
 XX acid; (c) a protein which elevates the efficiency of synthesis of
 XX isoprenoid compounds and comprises a 3 or 4 amino acid sequence
 XX optionally with 1 or more of the amino acids being deleted or
 XX substituted or an additional amino acid being inserted; (d) a protein
 XX which activates or catalyses the production of 2-C-methyl-D-erythreitol-
 XX 4-phosphate from 1-deoxy-D-xylulose-5-phosphate; or (e) a protein which
 XX activates a target compound or reaction and is a string end or hybrid of
 XX the DNA encoded in (a)-(e). Isoprenoid compounds are useful in drugs
 XX (e.g. for the treatment of heart diseases, osteoporosis and hemostasis,
 XX for preventing cancers and as immunopotentialators), health foods and
 XX antifouling coatings. The isoprenoid also inhibit enzymatic reactions
 XX on the non-nevralonate pathway and can be used as antibacterials and
 XX herbicides. The present sequence is used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 1182 BP; 185 A; 408 C; 417 G; 172 T; 0 other;
 SQ
 Query Match 100.0%; Score 1182; DB 20; Length 1182;
 Best Local Similarity 100.0%; Pred. No. 1.7e-187;
 Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 AACGGCCACGATCTCCGGTGGACAGCAGCACTCCGGTCTTTCAGGCCCTGGC 480
 DB 421 AACGGCCACGATCTCCGGTGGACAGCAGCACTCCGGTCTTTCAGGCCCTGGC 480
 QY 481 GGGCAGGACAGCGGCTGGCTGGAGCGGCTCAATCAACAGGGCTCCGGGCGGCTTCCG 540
 DB 481 GGGCAGGACAGCGGCTGGCTGGAGCGGCTCAATCAACAGGGCTCCGGGCGGCTTCCG 540
 QY 541 GACTGAGACCTTCAGAGCGCATCCGGCTGCACCTGGCGAGGGGAGGCCATCCCAAC 600
 DB 541 GACTGAGACCTTCAGAGCGCATCCGGCTGCACCTGGCGAGGGGAGGCCATCCCAAC 600
 QY 601 TGATCGAGAGCGCGGAATTTCTGGCTTCAGCGCGGACCGGATTCGAGGGCTGTCAT 720
 DB 601 TGATCGAGAGCGCGGAATTTCTGGCTTCAGCGCGGACCGGATTCGAGGGCTGTCAT 720
 QY 721 CCGCAATCATGTCATGCGATGATGGGCTTTCGCGAGGGGGCTGATGCGCATCTC 780
 DB 721 CCGCAATCATGTCATGCGATGATGGGCTTTCGCGAGGGGGCTGATGCGCATCTC 780
 QY 781 GGGCCCGCGACATGCGCCACGCGCATGCGATTCGCGTGAACCTGGCGGCTCGCGAG 840
 DB 781 GGGCCCGCGACATGCGCCACGCGCATGCGATTCGCGTGAACCTGGCGGCTCGCGAG 840
 QY 841 GTGCGCGCGCGCGGATGACCTGCGACAGATTCGCGAGGCTTCCAGAACCTCGAC 900
 DB 841 GTGCGCGCGCGCGGATGACCTGCGACAGATTCGCGAGGCTTCCAGAACCTCGAC 900
 QY 901 GAGGAGCGCTTCCGCGCTTGAGGCTTCCGCGAGATGCGGCGCGCGCTGTCG 960
 DB 901 GAGGAGCGCTTCCGCGCTTGAGGCTTCCGCGAGATGCGGCGCGCGCTGTCG 960
 QY 961 GGGCGCGCTTCAACGCGGCGCAAGAGATGCGCTGATCTTTCATGCGCGGAGCATC 1020
 DB 961 GGGCGCGCTTCAACGCGGCGCAAGAGATGCGCTGATCTTTCATGCGCGGAGCATC 1020
 QY 1021 GGGTTTCGACATGCGGCGGCTGCTGAGAGAGCGCTCGCGGCGCTTTCAGCGAGCC 1080
 DB 1021 GGGTTTCGACATGCGGCGGCTGCTGAGAGAGCGCTCGCGGCGCTTTCAGCGAGCC 1080
 QY 1081 CTGTTTCGAAAGTGCAGGCGCTTTCGAGGAGTCTGCGCATGAGCAATCTCTCGG 1140
 DB 1081 CTGTTTCGAAAGTGCAGGCGCTTTCGAGGAGTCTGCGCATGAGCAATCTCTCGG 1140
 QY 1141 AGAGCGGACAGAGAGCGCGGCTTCCGCGAGCAAGAAAG 1182
 DB 1141 AGAGCGGACAGAGAGCGCGGCTTCCGCGAGCAAGAAAG 1182

RESULT 2
 AAL40184
 ID AAL40184 standard; DNA; 1161 BP.
 XX
 XX AAL40184;
 XX
 XX 13-SEP-2002 (first entry)
 XX
 XX Isoprenoid related nucleic acid sequence SEQ ID No 96.
 XX
 XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;
 XX decaprenyl diphosphate synthase; ds.
 XX
 XX Sphingomonas trueperi.
 XX
 XX WO200226933-A2.
 XX
 XX 04-APR-2002.
 XX

CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
 CC activity or decaprenyl diphosphate synthase (Dds) activity, is useful for
 CC increasing production of Coq(10) in a cell having endogenous Dds
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.

XX Sequence 2017 BP; 278 A; 645 C; 722 G; 372 T; 0 other;

Query Match 32.2%; Score 380.4; DB 24; Length 2017;
 Best Local Similarity 60.3%; Pred. No. 1e-54;
 Matches 649; Conservative 0; Mismatches 421; Indels 6; Gaps 1;

QY 8 GCGTTCGATCTTGGGGCCACCGGCTCCATCGCGCAATCCACTTCGCTGATGC 67
 DB 585 GCGTTCGATCTTGGGGCCACCGGCTCCATCGCGCAATCCACTTCGCTGATGC 644
 QY 68 GGAAGGG 127
 DB 645 GAAA-----TCCGACGCTTCGAAAGTGTGGCGCTGACGCAAAATTGCGATGCGAGA 698
 QY 128 GACTGGCCGAATGCGCGCTGCGCTGAGAGCGGAGCTTCGCTACCGCGCATGAGACT 187
 DB 699 AGCTGGCTGCG 758
 QY 188 GCGTCCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247
 DB 759 GCGTCCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818
 QY 248 AGGCCATCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307
 DB 819 ATTGGGTGTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 878
 QY 308 CCGCGCGCGCTGCG 367
 DB 879 GCGGAGGCTCAAGCGCGGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 938
 QY 368 ACAAGGAAGCGCTGCTGACGCGAGGCGCAACTCTGATCGCGAGCGCGCGCGCG 427
 DB 939 ACAAGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 QY 428 CCAGATCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487
 DB 999 CGACGCTCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1058
 QY 488 ACAGCGCGCTGCG 547
 DB 1059 CGCCGAGGGGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1118
 QY 548 GCGTTCGAGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 607
 DB 1119 CGAAGGAAGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1178
 QY 608 TGGGCGACGAGATCTCTGACAGAGCGCGCTGATGTTCAACAAGCGCGCTGAGCTG 667
 DB 1179 TGGGCGCGCAAGATCTCTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1238
 QY 668 AGAGCGCGGAATCTTGGGCTTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCG 727
 DB 1239 AAGCTTCGACCTGTTCCCGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1298
 QY 728 CCATGTCGATGCGATGCGGCTTTCGCGAGCGGGGCGCGCGCGCGCGCGCGCG 787
 DB 1299 CCGTGTCTCATTCGATGCTGGAATATGTCGAGCGGATCGGTCTGGCGCGCGCG 1358
 QY 788 CCGACATCG 847
 DB 1359 CCGACATCG 1418
 QY 848 TCGCCCGGATCGACCTGCGACAGATTCGAGAGCTTCACCTTCAGAAAGCGTCAAGAGAAC 907
 DB 1419 GCCCGCGCGCTGACCTTCGACAGGAGGTAGCTGAGTTCGAAATTCGATCTGATTC 1478
 QY 908 GCTTCCGCGCGCTGAGGCTTTCGCGAGAGCTGATGCGCGCGCGCGCGCGCG 967

DB 1479 GCTTCCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1538
 QY 968 CCTTCAAGCG 1027
 DB 1539 TTCTCAATGCG 1598
 QY 1028 TGGACATGCG 1083
 DB 1599 TTGAATTCG 1654

RESULT 4

AA140185
 ID AA140185 standard; DNA; 1191 BP.

AC AA140185;

DT 13-SEP-2002 (first entry)

DE Isoprenoid related nucleic acid sequence SEQ ID No 105.

KW Isoprenoid; Coq(10); 1-deoxyxylulose-5-phosphate synthase; Dxs; Dds;

KM decaprenyl diphosphate synthase; ds.

OS Pseudomonas aeruginosa.

PN W0200226933-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30328.

PR 29-SEP-2000; 2000US-236580P.

PA (CRGT) CARGILL INC.

PI Gokarn R, Jessen H, Zidwick MJ;

PP WPI; 2002-416480/44.

PT Substantially pure polypeptides having e.g.,

PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the

PT production of isoprenoids, especially Coq(10)

PS Disclosure; Fig 31; 246pp; English.

CC The invention relates to methods and materials for the production of
 CC isoprenoids. More particularly the invention provides isolated nucleic
 CC acids, substantially pure polypeptides, host cells, and methods for
 CC producing various isoprenoid compounds. The polypeptides are useful for
 CC the production of isoprenoids, especially Coq(10). Expressing the pure
 CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
 CC activity or decaprenyl diphosphate synthase (Dds) activity, is useful for
 CC increasing production of Coq(10) in a cell having endogenous Dds
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.

CC Sequence 1191 BP; 176 A; 400 C; 413 G; 202 T; 0 other;

Query Match 24.2%; Score 286; DB 24; Length 1191;
 Best Local Similarity 55.8%; Pred. No. 4.7e-38;
 Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY 20 TTGGGCG 76
 DB 29 TCGGCG 88
 QY 77 GCGCGAGCG 136
 DB 89 GTTACGAAGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 148
 QY 137 AAATGGCG 196

Db 149 TCAGGACACGCCCCGCTATGCGGTGTGTCGCGAGACGCGCGCGGATTTGCTTTCAGG 208
 QY 197 CGCTGCGGAGAGCGCTGCGCGCGGAGCGGACCGAGGTCCGCGCGCGCGCGGAGCCATCG 256
 Db 209 GCTGCTGCGCGCGCGGATTCGCGACCGCGGCTGCTGTTCCGCGAGCGAGGCTGTGCG 268
 QY 257 CCGAGGCGCGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 316
 Db 269 AAGTGGCGCGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 328
 QY 317 TCGTGGCGCGGAAATGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGG 376
 Db 329 TCGCGCTCGACCTGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGG 388
 QY 377 GCGTGTGAGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 436
 Db 389 CGCTGGTGTGATTCGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGG 448
 QY 437 TCGCGGTGAGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 481
 Db 449 TCGCGGTGAGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 508
 QY 482 GCGAGGACACGCGCGCTGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGG 541
 Db 509 GCGTGGAGCGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 568
 QY 542 ACTGAGCGCTCGACGCGCATTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCG 601
 Db 569 AAGCGCGCGCTGAGCAACTTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCG 628
 QY 602 GGTCCATGGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 661
 Db 629 GGTCCATGGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 688
 QY 662 TGAATGAGACGCGCGCAATTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCG 721
 Db 689 TGAATGAGACGCGCGCAATTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCG 748
 QY 722 GCGCATTCATTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGCGGCGC 781
 Db 749 CCGAGCGCGTGTACCTGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGG 808
 QY 782 GCGCGCGCGACATGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 841
 Db 809 GCGATTCGCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 868
 QY 842 TSCCGCTGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 901
 Db 869 CCGCGCGTTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 928
 QY 902 AAGGACGCTTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 961
 Db 929 ACGACGCTTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 988
 QY 962 GCGCGCGCTTCGCGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 1021
 Db 989 CGGCGCATGCTGATGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 1048
 QY 1022 GGTTCCTGAGACATGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 1071
 Db 1049 GCTTCAGCGACATGCGCGCGCGGAGTGGACCATGTCGACATTCGCGCGCGCGCGGCGC 1098

RESULT 5

AA140187

AA140187 standard; DNA; 1161 BP.

AA140187;

13-SEP-2002 (first entry)

Isoprenoid related nucleic acid sequence SEQ ID No 107.

XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDXS;
 KM decaprenyl diphosphate synthase; ds.
 KW Streptomyces griseoliosporus.
 OS WO200226933-A2.
 PN 04-APR-2002.
 PD 28-SEP-2001; 2001WO-US30328.
 PF 29-SEP-2000; 2000US-236580P.
 PR (CRGI) CARGILL INC.
 PA Gokarn R, Jessen H, Zidwick MJ;
 PI WPI; 2002-416480/44.
 DR
 XX Substantially pure polypeptides having e.g.,
 PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
 PT production of isoprenoids, especially CoQ(10).
 XX Disclosures; Fig 31; 246pp; English.
 PS
 PS The invention relates to methods and materials for the production of
 CC isoprenoids. More particularly the invention provides isolated nucleic
 CC acids, substantially pure polypeptides, host cells, and methods for
 CC producing various isoprenoid compounds. The polypeptides are useful for
 CC the production of isoprenoids, especially CoQ(10). Expressing the pure
 CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)
 CC activity or decaprenyl diphosphate synthase (DPS) activity, is useful for
 CC increasing production of CoQ(10) in a cell having endogenous DXS
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.
 XX
 SO Sequence 1161 BP; 151 A; 399 C; 449 G; 162 T; 0 other;

Query Match 20.9%; Score 247.2; DB 24; Length 1161;
 Best Local Similarity 56.9%; Pred. No. 1.3e-32;
 Matches 474; Conservative 0; Mismatches 353; Indels 6; Gaps 1;

QY 231 GGTGCG 290
 Db 234 GGTGCG 293
 QY 291 GTGCG 350
 Db 294 GAAACGCGCATCACCGGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 353
 QY 351 CACGCTGCG 410
 Db 354 GGTGCTGTGCTGTGCGGT 413
 QY 411 GGGCGCGGAGAACG 470
 Db 414 GGGCG 467
 QY 471 GGGCG 530
 Db 468 GGGCG 527
 QY 531 GCGCTTCG 590
 Db 528 CCGCTTCG 587
 QY 591 CCATCCCACTGATGCG 650
 Db 588 GCAACCG 647
 QY 651 GGGCGCTGAGCTGATGAGACG 710

OY	705	CGGGCGGTCGTCATCCGCAATCCATGTCATGCAATGCGATGGGGCTTTCGACAGGGGG	764
Db	182654	CGAAGTCTCATTCATCCGCAATCCGTAATACACAGCATGGTGGCTTACCGCAGCGCTC	1826355
OY	765	CTGTATGGCCCATCTTCGGGCCCGCCGACATGCGCCAGCCATCGGATTCGCGTGAAC	824
Db	182634	CGTGTGGCGCAATCGGGGCATGCCGATATCCGATATCCGAAAGCCCATCGCTATTGTGGGTTT	182575
OY	825	GCCGGGTGCGGGCGAGGTGCCCGTCGCGCCGGATGACCTTGGACAGATTGCCAGCCTCAC	884
Db	182574	GCCCGAGGCGATTCATTCCTGGGTGTCGGCGGACCTTGATTTCGACGCATTTGCCGGCTGAC	182515
OY	885	CTTCAGAGAGCTGACGAGGAAGCTTTCCGGCCCTGAGGCTTGGCGAGACGTCATGGC	944
Db	182514	CTTCCAAAAGCCCGACTTTTGACCGCTTCCCTCGCTGAGGCTGGCTATTAAGGCATGAA	1824555
OY	945	GCGCGCGGCGCTGTTCGGGGCGCGCCCTTCAACGCGCGGCGCAAGAGATCGCGTCGATCTT	1004
Db	182454	CGCAGCGGGAGACCGCGGCCCTCGTATTTGAACGCCGCCAAGAAAGCCGCTCGCGGCTT	1823959
OY	1005	CATGCGCGGACGCAATCGGGTTTCTGGACATATGGCGGGCGGTGTCGAGGAGACGCTCGCGG	1064
Db	182394	TTTGGACGCGACAGATTAAATTATTCAGCAATATGCCAAAACGTGCGCCACTGTCTTGACA	1823359
OY	1065	CGTTTCGACGCGACCCCGCTGTGGAAAAGTC	1095
Db	182334	AGACTTTTCAGACGGCATATGGCGATATAGGG	182304

RESULT 8	
AA140188	
ID	AA140188 standard; DNA; 1185 BP.
XX	
AC	AA140188;
XX	
DT	13-SEP-2002 (first entry)
XX	
DE	Isoprenoid related nucleic acid sequence SEQ ID NO 108.
XX	
KW	Isoprenoid; Coq(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS; decaprenyl diphosphate synthase; ds.
XX	
OS	Neisseria meningitidis.
XX	
PN	MO200226933-A2.
XX	
PD	04-APR-2002.
XX	
PP	28-SEP-2001; 2001WO-0530328.
XX	
PR	29-SEP-2000; 2000US-236580P.
XX	
PA	(CRGI) CARGILL INC.
XX	
PI	Gokarn R, Jessen H, Zidwick MJ;
DR	WPI; 2002-416480/44.
XX	
PT	Substantially pure polypeptides having e.g.,
PT	1-deoxyxylulose-5-phosphate synthase activity, useful for the
PT	production of isoprenoids, especially Coq(10)
XX	
PS	Disclosure; Fig 31; 246pp; English.
XX	
CC	The invention relates to methods and materials for the production of
CC	isoprenoids. More particularly the invention provides isolated nucleic
CC	acids, substantially pure polypeptides, host cells, and methods for
CC	producing various isoprenoid compounds. The polypeptides are useful for
CC	the production of isoprenoids, especially Coq(10). Expressing the pure
CC	polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)
CC	activity or decaprenyl diphosphate synthase (DDS) activity, is useful for
CC	increasing production of Coq(10) in a cell having endogenous DDS
CC	activity. This polynucleotide represents a nucleic acid sequence relating

CC to the isoprenoid production of the invention.
XX
SQ Sequence 1185 BP; 255 A; 365 C; 326 G; 239 T; 0 other

Query Match	19.3%	Score 228.6;	DB 24;	Length 1185;
Best Local Similarity	53.3%	Pred. No. 1.5e-29;		
Matches 392;	Conservative	0;	Mismatches 489;	Indels 30;
				Gaps 4

Oy	9	CCCTGCATCTTTGGGGCCACACGGGTCATCTGGGCGAATCCACTTGCAGCTGTGATCG	68
Db	15	CTTACACATATTAAAGCAGTACCGGACACATAGCGGAAGACAGCTGAGCTGTCTCCC	73
Oy	69	GAAGGCGGGGCCGAGGCGCTTTCGCAACCGTCGTGACCGGGCGGCGCAATCCGGCG	128
Db	74	-----GGCACCCCGAAAAATTCCGCGGATTCCGGCTGGCAGGGGCAATAGCAGGTGAGAA	128
Oy	129	ACTGCGCGAAATGGCGGCTGCGCTGAAAGGCGAGGCTTGGCGCTACCCGCGCATGAGAGCTG	188
Db	129	ATTGAGCGGCTCAATGTCAAAAGTTTCCACCCCGGAATATGGCGTGTGGCGAGTGGCGAA	188
Oy	189	CTCTCCGCGCGTGGCGGAGCGCGTG-----GGCGGAGCGGGCACCGAGTGGCGGGG	242
Db	189	CGCCGCCCGGCTTGAAGCCCTGTTTGAACGCGGACGGCAGCGGCACTCAAGTTTTACAGG	248
Oy	243	GGCGCAGGC---CATGCGCGAGGCGCGCGACCGCGCGCGCACTGGACCATGTGCGCAT	299
Db	249	CGCGCAGGCAATTGGTTGACGTGTCTCTGTGCGCGACGAAGTCACGGGTGTATGGCGCAT	308
Oy	300	CGTGGGCGCGCGGCGCTCTGTGCCCGGAATGCGGGGCGCTGAAGCACGGCGCAGCTGGC	359
Db	309	CGTGGGCGCGTGGGCGTGTGCTTCCGCGCTCGCAGCGGGCGCAAAAAGGCAAAACATT	368
Oy	360	GCTCGCCACAAGAAAGAGCTCGTGAAGCGGAGGGCAACTCTGATGCGGACGGCGCAG	419
Db	369	TCTGGCGCAAGAAAGACGCTGGTGGTTTCGGCGCGTGTATTATGAAACCGCCGTC	428
Oy	420	GAACGGCGCCACGATCTGTGCGGTTGACAGCAGCAGCACTCCGGCTTTCAGGCGCTGGC	479
Db	429	AAACGGCGCGGCAATGCTGTGCGCGCTGCACAGCAACAGCGCGTTTTCAGATTTTGCC	488
Oy	480	GGGCGAGGACACGS-----CTGTGCTGAGCGCGCTCATCATACGGCGTC	524
Db	489	GCGGGAATTACACAGGTGCGCTGAACGAACAGGCACTGCGTTCGATTCTCAGCGCTTC	548
Oy	525	CGGCGGGCGGTCGCGCACTGGAAGCCTTCGAGCGCACTCCGCGTCGACCCGTGGCGAGG	584
Db	549	CGGGGGCGCGCTTCTGACCGCGCATTTAAACAGCTTGCAGACGATTAAGCCCGGACCAAG	608
Oy	585	GCAAGCCCATCCCACTGTTCATGGGCGCAGGGGATCTCCATGCAGAGCGCTCGATTT	644
Db	609	GGTCAAAACACCCAATATGGGCTATGGGAGCGCAAAATCTCGCTGATTCGCGCCACCATAT	668
Oy	645	CAACAAGGCGCTGAGCTGATTCAGAGACGGCGCAATCTTGGCGCTTTCGAGCGGACCGGAT	704
Db	669	GAACAAAGTTTGGAGCTGATTAAGCGGATTTGGCTTTCACTGTCCGCCCGCAAAACT	728
Oy	705	CGAGGCGGTGCTCCATCCGAATCCATTCGCACTGCGATGGGGCTTCTGAGAGGGGG	764
Db	729	CGAAGTGTCTATCATCTCCGAATCTGTGATACACACATATGGTGCCTACCGGACGGCTC	788
Oy	765	CGTGAATGCCCATCTCGGCGCGCGCGACATGCGCACGCCCATCGGATTCGGCGCTGAACTG	824
Db	789	CGTGTATGGCGCAACTGGGCAATCCCAATATGCAAGCGCATTCGCTTATTTGTTGGTTT	848
Oy	825	GCGCGGTGTGGGCGGAGGTGCGCCCTGCCCGGATTCAGACTTCGCACAGATTTGGAGGCTAC	884
Db	849	GCCCGACGCGATGATTCGGGTTGCGGCGACCTGGATTTTCAGCGCATTTGTCCGCGCTAC	908
Oy	885	CTTCCAGAGCTGACAGGAGAGCTTTTCGGCGCGCTGAGAGCTTGGCGGACGATGATGGC	944
Db	909	CTTCCAAAAGCCGCACTTTATACCGCTTCCCTGCTGAAGACTCGCCTATGAAAGCCATGA	968
Oy	945	GCGCGGGCGCTGTGGGCGCGCGCTTCAACGCGGCGCAAGAGATTCGCGCTGCATCATTT	1004

Db 969 CGCAGCGGAGCGCCGCGTATGACCCGCAAGCGCGCGCGCTT 1028
 OY 1005 CATGCCGAGCGATCGGTTCTGACATGCGGCGGTGAGGAGACCTCGCGG 1064
 Db 1029 TTGGACGACGATTAAATTACCGACATTCGCAAAACCGCGCCATTGTTTCA 1088
 OY 1065 GATTTCGACCGACCCCTGTTGCGAAAGTG 1095
 Db 1089 AGACTTTTCAGACGCGCATAGCGGACATAGCG 1119
 RESULT 9
 ABR83261
 ID ABR83261 standard; DNA; 1182 BP.
 AC ABR83261;
 XX 27-AUG-2002 (first entry)
 DE High growth methanotrophic bacterial strain DNA #32.
 XX High growth methanotrophic bacterial strain: Cl carbon substrate; gene:
 KW methane; methanol; Embden-Meyerhof carbon flux pathway; 16S rRNA; 3S;
 KW pyrophosphate dependent phosphotriokinase; nitrogen-containing compound;
 KW ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
 KW methane-containing environment; waste water treatment system; isoprenoid;
 KW nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
 XX Methylomonas 16a.
 OS WO200220728-A2.
 XX 14-MAR-2002.
 XX 28-AUG-2001; 2001WO-US26827.
 XX 01-SEP-2000; 2000US-229858P.
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 PA Koffas M, Odom JM, Schenkle A;
 PI WPI; 2002-452200/48.
 DR P-PSDB; ABG61582.
 XX New high growth methanotrophic bacterial strain, useful for producing
 PT single cell proteins, grows on a Cl carbon substrate, and comprises a
 PT functional gene encoding in Embden-Meyerhof carbon pathway
 XX Claim 29; Page 140; 157Pp; English.
 XX The invention relates to a high growth methanotrophic bacterial strain,
 CC which grows on a Cl carbon substrate e.g. methane and methanol, and
 CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
 CC gene coding a pyrophosphate dependent phosphotriokinase enzyme or a 16S
 CC rRNA. The bacterial strain is useful for the production of single cell
 CC protein and for the biotransformation of a nitrogen-containing compound,
 CC e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
 CC production of a feed product comprising a protein, carbohydrates and a
 CC pigment and for reducing oxygen demand, for removing nitrates and
 CC nitrates in methane-containing environments such as landfills, waste
 CC water treatment systems or anywhere that methane, oxygen and nitrates are
 CC present. The bacterial strain of the invention can be used as a
 CC denitrifying agent for the conversion of nitrate or nitrite to nitrous
 CC oxide with methane or methanol as a carbon source. It is also used in the
 CC production of biomass including proteins, carbohydrates and a wide
 CC variety of pigments (particularly for isoprenoid pigments for the
 CC purpose of generating animal feeds), in production of terpenoid and
 CC carotenoid compounds, useful as pigments and as monomers in polymeric
 CC materials and in production of exopolysaccharides at high levels.
 CC Sequences ABR83230-ABR83270 represent high growth methanotrophic
 CC bacterial strain DNA of the invention.

XX Sequence 1182 BP; 270 A; 301 C; 347 G; 264 T; 0 other;
 SQ Query Match 19.0%; Score 224.2; DB 24; Length 1182;
 Best Local Similarity 56.5%; Pred. No. 8.2e-29;
 Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;
 OY 277 GCCGACTGACCATCTGCGCCATCTGCGGCGGCGGCGCTGCGCGGATGCGGCG 336
 Db 283 GTGATACGCGTATGCGGCGCTATCGTCGCGCGCGCGGATGTTGCGGACCTTGGCGCG 342
 OY 337 CTGAAGCAGCGCGCGACGCTGCGCTCGCCAAAGAAAGAGCTGTGAGCGGCGCA 396
 Db 343 GCCAAGCGCGCAAAACCGCTGCTGTTGGCCAAAGAAAGAGCTGTGATGTCGGGACAA 402
 OY 397 CTCCTGATGCGGAGCGCGCCAGAGAAAGCGCGCAGATCTCGCGTGGACAGCGACAC 456
 Db 403 ATCTTCATGCGAGCGCGTCACGATTCGCGGCTGTGTCTCCGATGACAGCGAGCAC 462
 OY 457 TCCGCGCTTTTCAGCGCGCTGCGCGCGCGAGCACAGCGCGCTGC-----GTC 501
 Db 463 AAGCGCATCTTCAGTGCATGCGCGGCGGTTATAGCCAGCGCATACAGCCAAAGAGCGG 522
 OY 502 GAGCGGTCATATACAGCGCGCTCCGCGGCGCGCTTCGCGACTGAGCCTCGAGCGCATC 561
 Db 523 CGCGCGATTTTATTGACCGCTTCGCGTGGCCCATTTTCAGAGCAGCGCGATAGAAAGCTTG 582
 OY 562 CGCGCGTCGACCGGTGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGATC 621
 Db 583 TCCAGCGTACCGCGGATAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGATC 642
 OY 622 TCCATCGACAGCGCGCTCGATGTTCAACAGAGCGCTCGAGCTATCGAGCGCGCATTC 681
 Db 643 TCGGTCGATTTCCGCCACCATGATGAACAAAGGCTCGAATCGAAGCCTGCTGTTG 702
 OY 682 TTGCGCTTGAGCGCGGACCGGATCGAGCGGCTGTCATCCGATCATCGTCCATGCG 741
 Db 703 TTCAACATGAGCGATG 762
 OY 742 ATGTCGCTTTCG 801
 Db 763 ATGTCGCTTTCG 822
 OY 802 GCCATCGATTCG 861
 Db 823 CCGATGACCGAT 882
 OY 862 CTCGACAGATTCG 921
 Db 883 ATTTTCGAAGTAGGCG 942
 OY 922 AGCTTTCG 981
 Db 943 AGATTCGCTTATGAGCG 1002
 OY 982 AAGAGATTCG 1041
 Db 1003 AATGAATTCG 1062
 OY 1042 GTGTCGCGAG 1050
 Db 1063 ATCATCGAG 1071
 RESULT 10
 AAD35500
 ID AAD35500 standard; DNA; 1182 BP.
 XX AAD35500;
 AC AAD35500;
 XX 25-JUL-2002 (first entry)
 DE Methylomonas 16a sp. D-1-deoxyxylose-5-phosphate reductoisomerase gene.

PF 29-AUG-2001; 2001WO-US26852.
 XX
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cheng Q, Kofas M, Norton KC, Odom JM, Picatagallo SK, Rouviere PE;
 PI Schenzle A, Tomb J;
 XX
 DR WPI: 2002-383051/41.
 DR P-PSDB; AA080326.
 XX
 PT Novel nucleic acid molecule encoding a isoprenoid biosynthetic enzyme,
 PT isolated from *Methylobionas* 16a, useful for the production of isoprenoid
 PT compounds -
 XX
 PS Claim 2; Page 68; 84pp; English.
 XX
 CC The present invention relates to a new nucleic acid molecule encoding
 CC an isoprenoid biosynthetic enzyme isolated from *Methylobionas* 16a.
 CC The invention is useful for obtaining a nucleic acid molecule
 CC encoding an isoprenoid compound biosynthetic enzyme, and for the
 CC microbial production of isoprenoid compounds. The molecules of the
 CC invention are also useful for regulating isoprenoid biosynthesis in an
 CC organism and for producing recombinant organisms for producing various
 CC isoprenoid compounds. The nucleic acid is also useful for feed additive,
 CC for the production of keratensoids and their derivatives, isoprenoid
 CC intermediates, and as pure products useful as pigments, flavours and
 CC fragrances. The present nucleic acid sequence encodes the *Methylobionas*
 CC 16a open reading frame 2 (ORF2) dxr (1-deoxyxylulose-5-phosphate
 CC reductoisomerase enzyme) protein of the invention, as described above.
 XX
 SQ Sequence 1182 BP; 270 A; 301 C; 347 G; 264 T; 0 other;

Query Match 19.0%; Score 224.2; DB 24; Length 1182;
 Best Local Similarity 56.5%; Pred. No. 8.2e-29;
 Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

277 GCCGACGTGACACATGTCGGCCATGTCGGCGCGCGCGCTGCTGCCCGGATGCGGCG 336
 1
 283 GTGATACGATGATGCGGTATGCGTATGCGCGCGCGGATGTTGTCGACCTTGCGCGC 342
 1
 337 CTGAGACAGCGCGCGACGCTGCGCTGCGCCACAGAGAAAGCTTCGTGACGCGAGCGCAA 396
 1
 343 GCCAAGCGCGCGCAAAACCGCTGCTGGCCACAGAGAAAGCTTCGTGATGATGCGGACAA 402
 1
 397 CTCCTGATGCGGCG 456
 1
 403 ATCTTATGATGCG 462
 1
 457 TCCGCGGTCTTTCAGCG 501
 1
 463 AACGCCATCTTTCATGTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 522
 1
 502 GAGCGCGGTATCATCATGCG 561
 1
 523 CGCGCGATTTTATGACCG 582
 1
 562 CGCGCGTGCACCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGATC 621
 1
 583 TCCAGGTCATGACCGAT 642
 1
 622 TCCATGACAGCGAT 681
 1
 643 TCGGTGATTCGAT 702
 1
 682 TTGCGGTGAGCG 741
 1
 703 TTCAACATGAGCG 762
 1
 742 ATGCTGGGCTTTCG 801
 1
 763 ATGCTGGACTATGTCATGCTGCTGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
 1

QY 802 GCCATCGGATTCG 861
 1
 DB 823 CCATATGCGAT 882
 1
 QY 862 CTCGACAGATTCG 921
 1
 DB 883 ATTTGCAAGTATGAGCG 942
 1
 QY 922 AGGCTTCG 981
 1
 DB 943 AGATTCGCTTATGAGCG 1002
 1
 QY 982 AAGGATTCG 1041
 1
 DB 1003 AATGAATTCGCTGCG 1062
 1
 QY 1042 GTGCTCGAG 1050
 1
 DB 1063 ATCATCGAG 1071
 1

RESULT 12
 ID ABL53861 standard; DNA; 11575 BP.
 XX
 AC ABL53861;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Plasmid pTUS75:dxs:dxr:Tn5kn.
 XX
 KM Plasmid pTUS75:dxs:dxr:Tn5kn; vector; limonene synthase;
 KM monoterpene; cyclic terpene; ds.
 XX
 OS Chimeric - *Methylobionas* sp.
 OS Chimeric - Unidentified.
 XX
 PN WO200220815-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 28-AUG-2001; 2001WO-US26853.
 XX
 PR 01-SEP-2000; 2000US-229858P.
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Dicosimo DJ, Kofas M, Wang S;
 PI WPI: 2002-339805/37.
 XX
 PS Example 6; Page 53-60; 63pp; English.
 XX
 CC The present sequence is that of vector plasmid pTUS75:dxs:dxr:Tn5kn.
 CC This is a derivative of RK2, a broad host-range plasmid, modified to
 CC include kanamycin resistance and 2 *Methylobionas* genes: dxs encoding
 CC 5'-deoxy-D-xylulose synthase; and dxr encoding 5'-deoxy-D-xylulose
 CC reductoisomerase. A *Mentha spicata* limonene synthase gene (see
 CC ABL53856) encoding a truncated enzyme (see ABL75749) was ligated
 CC into the vector, creating plasmid pDH3, in which the limonene
 CC synthase gene was flanked on the 5' end by dxs and on the 3' end by
 CC dxr. pDH3 was transferred to *Methylobionas* sp. 16a by triparental
 CC conjugal mating. Cells that received the limonene synthase gene
 CC demonstrated to produce limonene. This is an example of the method
 CC of the invention, in which a C1 metabolizing host cell is
 CC transformed with a gene encoding a cyclic terpene synthase, and is

CC used to produce a monoterpane.

XX Sequence 11575 BP; 2395 A; 3273 C; 3548 G; 2359 T; 0 other;

Query Match 19.0%; Score 224.2; DB 24; Length 11575;

Best Local Similarity 56.3%; Pred. No. 6.9e-29;

Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

```

QY 277 GCCGACGTGAGACATGTCGGCCATGTCGGGCGCCGCGGCGCCCTGCGCCGGAATGCGGGCG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3339 GTGATACGGTATGATGCGGCTATGTCGCGCGCGCCGATTTGTTGCGACCTTGCGCGCG 3398
QY 337 CTGAAGACGCGCCGACGCTGCGCCCTGCCAACAAGAAAGCTCTGTACGCGAGCGCA 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3399 GCCAAGCGCGGCAAAACGCTGCTGTGGCCAAAGAAGGCTTGTGTATGTCGGGACAA 3458
QY 397 CTCTGTATGCGGAGCGCCGACAGAAAGCGCGCCACGATCCCTGCGGCTGTGACAGGAGCAC 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3459 ATCTTCATGCAAGCGCGCTGACGATTCGCGCGCTGTGTGCTGCGCATAGACAGGAGCAC 3518
QY 457 TCCGCGCTCTTTCAGGCGCTGCGGCGGAGAGACAGGCGCTC-----GTC 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3519 AAGCCCATCTTTCAGTGCATGCCGCGGCTTATACGCCACGCTACAGCCAAACAGCGCG 3578
QY 502 GAGCGCTCATCATACGCGCTGCGGCGCGCTTCCGCGCATGAGCCTTCGAGCGCATC 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3579 CGCGCGATTTTATGACGCGCTTCCGCTGCGCCATTTGACGAGCGCGCATAGAAACGTTG 3638
QY 562 CGCGCGTACCGCGGCGGCGGCGGCGCGCGCCATCCCAACTGTGCTCATGCGGCGAGCATC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3639 TCCAGCGTACGCGCGGATCAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3698
QY 622 TCCATGCAAGACGCGCTGATGTCACAAAGCGCGCTGAGATGAGACGCGGAAATTC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3699 TCGGTCGATTCGCGCAGCATGATGACAAAGGTCGCAACTGATGCAAGCGCTGTTGTG 3758
QY 682 TTGCGCTTCGAGCGCGGACCGGATGAGCGCGTCCATCCGCAATCATGTCATCGC 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3759 TTCAACATGAGCGCGGACCAATGAGTCTCATTTCCGAGAGCATCATTCATTCG 3818
QY 742 ATGATGCGCTTCTCGAGCGGCGGCTGATGCGCCATCTCGGCGCGCGGACATGCGCGCAC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3819 ATGATGCGCTTCTCGAGCGGCGGCTGATGCGCCATCTCGGCGCGCGGACATGCGCGCAC 3878
QY 802 GCCATCGGATTCGCGCTGAACTGCGCGGCTGCGGCGGAGTGCCTGCGCGGATCGAC 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3879 CCGATACGCGCGCGGATGCGGCTGCGCGGAGCGCTTTGACTGCTGCTGCGCGCGCGTGGAT 3938
QY 862 CTCGACAGATGCGGAGCTTACCTTCAGAAAGCTGACGAGAAAGCTTTCCGCGCTG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3939 ATTTTCGAAGTAGGCGCAAGTGAATTTGCAAAACCGCACTTGAAACGGTTTCTTGTG 3998
QY 922 AGGCTTCGCGAGAGCTGATGCGCGCGCGCGCTGTCGCGCGCGCGCTTCAACGCGCGC 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3999 AGATTCGCTTATGAAGCATCAAGTGTGTGATATGCAACGCGTATGAGACGCGC 4058
QY 982 AAGGAGATCGCGCTGATCATTCATCGCGCGAGCATCGGCTTCTGTGACATGCGCGCG 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4059 AATGAATTTGCTGTGCAAGCGTTTAAATGAAGAAGTCAATTCACATGACATCGCGGCTC 4118
QY 1042 GTGCTCGAG 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4119 ATCATCGAG 4127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

AA253883 standard; DNA; 888 BP.

XX AA253883;

XX 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 572 partial DNA sequence SEQ ID NO:1715.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

XX antibacterial; gene therapy; ds.

XX Neisseria meningitidis.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.

XX 01-MAY-1998; 98US-0083758.

XX 31-JUL-1998; 98US-0094869.

XX 02-SEP-1998; 98US-0098994.

XX 02-SEP-1998; 98US-0099062.

XX 09-OCT-1998; 98US-0103749.

XX 09-OCT-1998; 98US-0103794.

XX 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masiagnani V, Mora M;

XX Petersen J, Piza M, Rappulli R, Ratti G, Scalato E, Scarselli M;

XX Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX P-PSDB; AAY75121.

XX Novel Neisserial polypeptides predicted to be useful antigens for

XX vaccines and diagnostics

XX Claim 7; Page 879; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

XX and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent

XX PCR primers used in the exemplification of the present invention. The

XX polypeptides, the polynucleotides, antibodies and compositions of

XX the invention can be used as vaccines, as diagnostic reagents, and as

XX immunogenic compositions. The polypeptides can be used in the

XX manufacture of medicaments for treating or preventing infection due to

XX Neisserial bacteria (e.g. meningitis and septicemia), to detect the

XX presence of Neisseria bacteria, or to raise antibodies. They may also

XX be used to screen for agonists or antagonists, which may themselves

XX have use as antibacterial agents. The polynucleotides of the invention

XX may also be used in gene therapy protocols.

XX Sequence 888 BP; 186 A; 280 C; 242 G; 180 T; 0 other;

Query Match 18.4%; Score 217.2; DB 21; Length 888;

Best Local Similarity 55.2%; Pred. No. 1.2e-27;

Matches 454; Conservative 0; Mismatches 353; Indels 15; Gaps 1;

```

QY 289 ATGTCGGCATGTCGGCGCGCGGCTCGTCCGGAATGCGGCGCTGAGACGCGC 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGTCGGCATGTCGGCGCGGCTCGTCCGGAATGCGGCGCTGAGACGCGCGCAAAAAGCG 60
QY 349 CGCAGCGTGGCGCTGCGCAACAGAAAGCTGTGACGCGGAGGCAATCTCTATGCGG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AAACCATTTTATGTGGCAAAACAAGAAACGCTGTGTGTTCCGCGGCTTGTATGGA 120
QY 409 ACGGCCAGAGAGAGCGCGCCACGATCTGCGGCTGAGACAGGAGCTCCGCGCTTT 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ACCGCCGTCGCAACGCGCGGCGAGTGTGCTCCGTCACACAGCAACACGCGCTTTC 180
QY 469 CAGCGCTGCGGCGGCGAGAGACAGG-----CCTGCTGACGCGCGTCAATC 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CAACTTTTGGCGCGGATTCGCGCGCGCTGTGAAACAGACAGGCGCTTCATATC 240

```

QY 514 ATCAGCGGCTCGCGGCGGCTTCCGACGTGAGGCTCGAGGCGATCCGCGCTGCACC 573
 DB 241 CTGACCGGCTTCGGCGCGCGCTTCTGACCGCGGATTTAAACAGCTTCAGCCGATTAAG 300
 QY 574 GTGGCCGAGCGGAGGCCATCCCACTGTGTCATGGCCAGGATTCATCGACAGC 633
 DB 301 CCGCGCGGCTCGCAAAAGCCCAATTCGCGTATGGGCGCAAAATCTCCGCTGATTC 360
 QY 634 GCGTCATGTCACAAAGGCGCTCGAGCTGATGAGAGCGGCGCAATTCCTCGGCTCGAG 693
 DB 361 GCCACCATATGACAAAGTTGGAGCTGATTTGAAGCGCATTTGGCTGTTCAACTGCC 420
 QY 694 CCGGAGCGGAGGCGGCTGCTGCTCCGCAATCCATGCTCCATGCTGATGGGCTTC 753
 DB 421 CCGGCAAACTGAGAGTCGATCAGATCCGCAATCCGATGATACAGCATGGGCGCTAC 480
 QY 754 TGGAGCGGCGGCTGATGCGCCATCTCGGCGCGCGCGCAATGGCGCAACGCGCATGATTC 813
 DB 481 CCGGAGCGGCTCGCTGCTGCGCAACCTGCGCAATCCGATATGCGCAACGCCCATCGCTTAT 540
 QY 814 GCGCTGACCTGCGCGGCGGCTGCGGCGAGGTGCGCGCGGATTCGACCTCGCACAGATT 873
 DB 541 TGTTCGGCTTTCGCGAGCGCATGCTGCGGTGCGCGACCTGATTCGACGCAATTG 600
 QY 874 GCGAGCTCAGCTTCGAGAGCCTGACGAGAGCCTTTCGCGCGCTTCGAGGCTTCGCGCA 933
 DB 601 TCGCGCTGACCTTCCAAAAGCGCGACTTTCGACGCTTCCTCCGCTGAGGCGTCCGCTAT 660
 QY 934 GACGTATGCGGCGGCGCGGCTGCTGCGGCGCGCTTCACAGCGCGCGCAAGAGATCGC 993
 DB 661 GAAGCATATGACGACGCGAGCGCGCGCTGCTGATGACGCGCGCGCAAGCGCGC 720
 QY 994 CTGATCATTTATTCGCGCGAGCATCGGCTTCGAGCATGGCGGCGGCTGCGAGAG 1053
 DB 721 GTGCGCGCTTTCGAGCGACGATTAAGTTACGACATTCGCAAAACGTCGCGCAC 780
 QY 1054 ACCTGCGGCGGCTTCGACGCGACCGCGCTGTCGAGAAAGTG 1095
 DB 781 TGTCTTGCAAGACTTTTCAGAGCGCATAGGCGGATATAGG 822
 RESULT 14
 AAH52058
 ID AAH52058 standard; DNA; 1311 BP.
 AC AAH52058;
 XX 04-SEP-2001 (first entry)
 DT Mycobacterium tuberculosis potential drug target gene SEQ ID 112.
 DE Mycobacterium tuberculosis potential drug target gene SEQ ID 112.
 XX Drug target: growth; organism viability; characterisation; ds.
 KW Mycobacterium tuberculosis.
 OS Mycobacterium tuberculosis.
 XX WO200135317-A1.
 PN 17-MAY-2001.
 XX 13-NOV-2000; 2000WO-US31152.
 PF 12-NOV-1999; 99US-0165086.
 PR 12-NOV-1999; 99US-0165124.
 PR 01-FEB-2000; 2000US-0179531.
 XX (REGC) UNIV CALIFORNIA.
 PA Eisenberg D, Rotstein SH, Marcotte EM;
 PI MPI: 2001-329193/34.
 DR P-PSDB; MAG81207.
 XX

PT Identifying nucleotide or polypeptide sequence for use as drug target,
 PT involves providing algorithm that analyzes a functional relationship
 PT between nucleotide or polypeptide sequences, and comparing the
 PT sequences
 XX Disclosure; Page 125-126; 207pp; English.
 PS
 XX This invention relates to a method for identifying a nucleotide or
 CC polypeptide sequence that may be a drug target, or essential for growth
 CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
 CC represent DNA encoding proteins MAG81096 - MAG81241, Mycobacterium
 CC tuberculosis proteins which are potential drug targets. The DNA and
 CC protein sequences are used to illustrate the method of the invention. The
 CC method involves providing an unknown nucleotide or polypeptide sequences,
 CC and comparing it to a number of sequences along with at least one
 CC algorithm capable of analysing a functional relationship between
 CC nucleotide and polypeptide sequences. The method is useful for
 CC characterising the function of nucleic acids and polypeptides that may be
 CC useful as a target for a drug or essential for the growth or viability of
 CC an organism.
 XX
 SO Sequence 1311 BP; 201 A; 429 C; 461 G; 220 T; 0 other;
 Query Match 18.2%; Score 215.6; DB 22; Length 1311;
 Best Local Similarity 54.2%; Pred. No. 2.2e-27;
 Matches 462; Conservative 0; Mismatches 384; Indels 6; Gaps 1;
 QY 201 GCGCGAGGCGGCTGCGCGGAGCGGCGACGAGTGCAGGCGGCGGCGCGCGACGCGCA 260
 DB 282 GCAGGCGGCGGCGGCGGCGGCGACATCCCTACAGGATTCGAGCGCGCGCGGCGCT 341
 QY 261 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 320
 DB 342 GGTGAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 401
 QY 321 GCGCGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 380
 DB 402 ACCGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 461
 QY 381 CGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 440
 DB 462 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 515
 QY 441 GGTGAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 500
 DB 516 GGTGAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 575
 QY 501 CGAGCGGCTGATCATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 560
 DB 576 CGCGAGAGGCTGCTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 635
 QY 561 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 620
 DB 636 CGAGCATGTCACCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 695
 QY 621 CTCATGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 680
 DB 696 CAGCGTGAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 755
 QY 681 CTTGCGCTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 740
 DB 756 GTTCGCGCATTCCTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815
 QY 741 GATGCGGCGGCTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 800
 DB 816 GATGTCACCTTCATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
 QY 801 CCGCATGAGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 860
 DB 876 ACCGATTTGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 935
 QY 861 CTTGCGACAGATTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 920

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2003, 00:00:39 ; Search time 3441 Seconds

(without alignments)
9996.953 Million cell updates/sec

Title: US-09-673-198-31

Perfect score: 1182

Sequence: 1 atgcgcagccgtcgtacctt.....gtctccgcagcagaagaag 1182

Scoring table:

IDENTITY-NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hlg_hum:*
- 31: em_hlg_inv:*
- 32: em_hlg_other:*
- 33: em_hlg_mus:*
- 34: em_hlg_pln:*
- 35: em_hlg_rtd:*
- 36: em_hlg_mam:*
- 37: em_hlg_vrt:*
- 38: em_sy:*
- 39: em_hlg_hum:*
- 40: em_hlg_mus:*
- 41: em_hlg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1182	100.0	1182	6	E59203	E59203 Method for
2	1182	100.0	1182	6	E64384	E64384 Process of
3	480.6	40.7	333800	1	SME591792	AL591792 Sinorhizo
4	434.2	36.7	9436	1	AE008174	AE008174 Agrobacte
5	434.2	36.7	10301	1	AE009208	AE009208 Agrobacte
6	352.6	29.8	12977	1	AE005865	AE005865 Caulobact
7	305.4	25.8	11150	1	AE012791	AE012791 Chlorobiu
8	292.2	24.7	201050	1	AL646064	AL646064 Ralstonia
9	290.6	24.6	10029	1	AE012236	AE012236 Xanthomon
10	286	24.2	1191	1	AF282879	AF282879 Pseudomon
11	286	24.2	14537	1	AE004785	AE004785 Streptomy
12	278.6	23.6	39744	1	SC5H4	AL355913 Streptomy
13	255.2	21.6	10242	1	AE011773	AE011773 Xanthomon
14	247.2	20.9	1161	1	AB049187	AB049187 Streptomy
15	238.2	20.2	13538	1	AE002375	AE002375 Neisseria
16	238.2	20.2	349980	6	AX043922	AX043922 Sequence
17	228.6	19.3	340806	1	NMA122491	AL162752 Neisseria
18	224.2	19.0	1182	6	AX394964	AX394964 Sequence
19	224.2	19.0	1182	6	AX398486	AX398486 Sequence
20	224.2	19.0	1182	6	AX398794	AX398794 Sequence
21	224.2	19.0	11575	6	AX398630	AX398630 Sequence
22	215.6	18.2	39991	1	MTCY274	Z74024 Mycobacteri
23	214.6	18.2	14715	1	AE007117	AE007117 Mycobacte
24	211	17.9	299950	1	AP005372	AP005372 Thermosyn
25	201.2	17.0	34083	1	AF124757	AF124757 Mycobacter
26	196.4	16.6	1167	1	ZMO250714	AJ250714 Zygomonas
27	195.4	16.5	10487	1	AE003942	AE003942 Xylella
28	183	15.5	11576	1	AE001994	AE001994 Deinococc
29	181	15.3	318200	1	MEPRN16	AL583922 Mycobacte
30	179.2	15.2	1115	6	AX432668	AX432668 Sequence
31	176.2	14.9	36985	1	MSG81529CS	L76824 Mycobacteri
32	172.4	14.6	5476	1	SLE250721	AJ250721 Synechoco
33	166	14.0	208780	1	BSUB0009	Z99112 Bacillus su
34	152.6	12.9	146271	1	SYCSLRB	D64000 Synchocyst
35	151.8	12.8	303249	1	AP001515	AP001515 Bacillus
36	146.6	12.4	250050	1	AL591978	AL591978 Listeria
37	145.8	12.3	337200	6	AX122293	AX122293 Sequence
38	145.8	12.3	337200	6	AP005280	AP005280 Coryneb
39	145.8	12.3	349880	6	AX127149	AX127149 Sequence
40	145.8	12.3	349880	6	AX127150	AX127150 Sequence
41	142.8	12.1	22997	1	AE008705	AE008705 Salmonell
42	141.2	11.9	268050	1	AL627266	AL627266 Salmonell
43	139.6	11.8	10808	1	AE001235	AE001235 Treponema
44	139.6	11.8	348550	1	AP003596	AP003596 Nostoc sp
45	132.8	11.2	1706	8	AF367205	AF367205 Oryza sat

ALIGNMENTS

RESULT 1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

[illegible]

D	b661	CTGATCGAGACGGCGGAAATTCTTGCGCTTCAGACCGGACGGCATGTGAGCGGTGCAT	720
O	y721	CCGCATCCATCCTGTCATGCGATGGTGGGCTTTCTGCCACGGGGGCTGTATGGCCATCTC	780
D	b721	CCGCATCCATCCTGTCATGCGATGGTGGGCTTTCTGCCACGGGGGCTGTATGGCCATCTC	780
O	y781	GACCCCGCCGACATGCGCACGACCATCGGAATTCGGCGTAGACTGGCGGATGCGGGGAG	840
D	b781	GACCCCGCCGACATGCGCACGACCATCGGAATTCGGCGTAGACTGGCGGATGCGGGGAG	840
O	y901	GTCGCCCTGTCGCCCGGATCGACCTTCGACAGATTTGCGAAGCTTCACCTTCACAGAGCTTAC	900
D	b901	GTCGCCCTGTCGCCCGGATCGACCTTCGACAGATTTGCGAAGCTTCACCTTCACAGAGCTTAC	900
O	y961	GAGCAAGCTTTCCGGCGCCCTGAGGCTTCGGGAGACGTCATGGGGGCGCGGGCTGTGCG	960
D	b961	GAGCAAGCTTTCCGGCGCCCTGAGGCTTCGGGAGACGTCATGGGGGCGCGGGCTGTGCG	960
O	y961	GCGCGCGCTTCACGCGCGCCAGAAGATCGCGCTCATTTTCATTCGCGCGACGATC	1020
D	b961	GCGCGCGCTTCACGCGCGCCAGAAGATCGCGCTCATTTTCATTCGCGCGACGATC	1020
O	y1021	GGGTTTCTGACATGCGCGCGGTGTGAGGAGACGTCGCGGGCGCTTTCGACCGACCCC	1080
D	b1021	GGGTTTCTGACATGCGCGCGGTGTGAGGAGACGTCGCGGGCGCTTTCGACCGACCCC	1080
O	y1081	CTGTTCCGAAAAAGTCCCGACGCGCTTTCAGGAAGTGTGCGCATGACATCTGCTCGG	1140
D	b1081	CTGTTCCGAAAAAGTCCCGACGCGCTTTCAGGAAGTGTGCGCATGACATCTGCTCGG	1140
O	y1141	AGAGCGGCAAGAGAACGCCCGGCTTCGCCACGACAAAAGG	1182
D	b1141	AGAGCGGCAAGAGAACGCCCGGCTTCGCCACGACAAAAGG	1182
R	SULT 2		
L	E64384	1182 bp DNA linear PAT 31-JAN-2002	
D	LOCUS	E64384	
I	DEFINITION	Processes of microbial production of isoprenoid compound.	
A	VERSION	E64384.1 GI:18622923	
K	KEYWORDS	JP 2000300256-A/23.	
S	SOURCE	Rhodobacter sphaeroides.	
O	ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Rhodobacter.	
R	REFERENCE	1 (bases 1 to 1182)	
A	AUTHORS	Miyake,K., Hashimoto,S., Motoyama,H., Ozaki,A., Set,H., Kuzuyama,T. and Takahashi,S.	
T	TITLE	Process of microbial production of isoprenoid compound	
J	JOURNAL	Patent: JP 2000300256-A 23 31-OCT-2000;	
C	COMMENT	OTOWA HAKKO KOGYO CO LTD OS Rhodobacter sphaeroides PN JP 2000300256-A/23 PD 31-OCT-2000 PF 12-APR-1999 JP 1999104589	
P	PI	KOICHIRO MIYAKE,SHINICHI HASHIMOTO,HIROAKI MOTUYAMA,AKIO OZAKI,	
P	PI	HARUO SETO,	
P	PI	TOMOHISA KUZUYAMA,SHUNJI TAKAHASHI	
P	PC	C12M15/09,C12P7/04,C12P13/00,C12P21/02/(C12P7/04, PC C12R1/19),	
P	PC	(C12P7/04,C12R1/18),(C12P7/04,C12R1/01),(C12P7/26,C12R1/19),	
P	PC	(C12P7/26,C12R1/18),(C12P7/26,C12R1/01),(C12P13/00,C12R1/19),	
P	PC	(C12P13/00,C12R1/18),(C12P13/00,C12R1/01),(C12P21/02,C12R1/19),	
P	PC	(C12P21/02,C12R1/18),(C12P21/02,C12R1/01),C12M15/00 CC	
F	FT	Key Location/Qualifiers	
F	FT	CDS (1)..(1182).	
S	SOURCE	Location/Qualifiers I..1182	

/organism="Rhodobacter sphaeroides"
/db_xref="taxon:1063"
BASE COUNT 185 a 408 c 417 g 172 t
ORIGIN

Query Match 100.0%; Score 1182; DB 6; Length 1182;
Best Local Similarity 100.0%; Pred. No. 1,1e-123;
Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGAGGCTGTCATCTTTGGGGCCACCAGGCTCCATGGGCAATCCACTTCGACCTC 60
DB 1 ATGCGAGGCTGTCATCTTTGGGGCCACCAGGCTCCATGGGCAATCCACTTCGACCTC 60
QY 61 GTCATCGGAGAGGG 120
DB 61 GTCATCGGAGAGGG 120
QY 121 ATCCGGGAGCTGCGCCAAATGGCGCTGCGCTGAAAGGCGAGCTTCGCTCACCGGCAT 180
DB 121 ATCCGGGAGCTGCGCCAAATGGCGCTGCGCTGAAAGGCGAGCTTCGCTCACCGGCAT 180
QY 181 GAGGAGTGGCTGCGCGCGCTGCGCGAGGCGCTGCGCGAGGCGGAGCGAGGCTGCGGGC 240
DB 181 GAGGAGTGGCTGCGCGCGCTGCGCGAGGCGCTGCGCGAGGCGGAGCGAGGCTGCGGGC 240
QY 241 GGGGCGCAGGCGCATCGCGAGGCGCGCGAGGCGCGCGAGGCGCGAGGCGCGAGGCGCGATC 300
DB 241 GGGGCGCAGGCGCATCGCGAGGCGCGCGAGGCGCGCGAGGCGCGAGGCGCGAGGCGCGATC 300
QY 301 GTGGGCGCGCGCGCGCTGCTGCGCGAGGCGCGCTGAAAGCAGCGCGCGCGCGCGCGCG 360
DB 301 GTGGGCGCGCGCGCGCTGCTGCGCGAGGCGCGCTGAAAGCAGCGCGCGCGCGCGCGCG 360
QY 361 CTCGGCAACAGGAAGGCTGCTGAGGAGGCGGCACTCTGATGGGAGCGCGCGCGCGAGG 420
DB 361 CTCGGCAACAGGAAGGCTGCTGAGGAGGCGGCACTCTGATGGGAGCGCGCGCGCGAGG 420
QY 421 AACGGCGCGCGAGGCTGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCG 480
DB 421 AACGGCGCGCGAGGCTGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCG 480
QY 481 GGGGAGGAGACGCGCTGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCG 540
DB 481 GGGGAGGAGACGCGCTGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCG 540
QY 541 GACTGAGGCTGAGGCGCATCGCGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGATC 600
DB 541 GACTGAGGCTGAGGCGCATCGCGCTGCGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGATC 600
QY 601 TGGTCCATGGGCGAGGCGATCTCCATCGACAGCGCGCTCGATGTTCAACAAGGCGCTCGAG 660
DB 601 TGGTCCATGGGCGAGGCGATCTCCATCGACAGCGCGCTCGATGTTCAACAAGGCGCTCGAG 660
QY 661 CTGATGAGACGCGCGCAATCTTCGCTGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGATC 720
DB 661 CTGATGAGACGCGCGCAATCTTCGCTGCGAGGCGCGAGGCGCGAGGCGCGAGGCGCGATC 720
QY 721 CCGCAATCATGCTGCTCATGAGATGTTGGGGCTTGGAGGCGGGGCGCGATGCGCATCTC 780
DB 721 CCGCAATCATGCTGCTCATGAGATGTTGGGGCTTGGAGGCGGGGCGCGATGCGCATCTC 780
QY 781 GGGCGCGCGAGCATGGCGCACGCGCATCGATCGGATTCGCGTGAACGGCGGGTCCGGCGAG 840
DB 781 GGGCGCGCGAGCATGGCGCACGCGCATCGGATTCGCGTGAACGGCGGGTCCGGCGAG 840
QY 841 GTGCGCGCGCGCGAGTGCATCGCGACAGATTTGCGAGCTTCACCTTCCAGAAAGCTGAC 900
DB 841 GTGCGCGCGCGCGAGTGCATCGCGACAGATTTGCGAGCTTCACCTTCCAGAAAGCTGAC 900
QY 901 GAGGAACGCTTTCGCGCGCTGAGGCTTGGCGGAGAGGCTCATGGCGGCGCGCGCTGTCG 960
DB 901 GAGGAACGCTTTCGCGCGCTGAGGCTTGGCGGAGAGGCTCATGGCGGCGCGCGCTGTCG 960

QY 961 GGGCGCGCTTCAACGCGGCGAGAGATCGCGCTGATCATTTTCATCGCGGAGCATC 1020
DB 961 GGGCGCGCTTCAACGCGGCGAGAGATCGCGCTGATCATTTTCATCGCGGAGCATC 1020
QY 1021 GGGTTTCTGACATGGCGCGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 GGGTTTCTGACATGGCGCGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 CTGTTGGGAAGATGGCGCGAGGCGCTTGGAGAGAGTCTGCGCATTCGCTCGG 1140
DB 1081 CTGTTGGGAAGATGGCGCGAGGCGCTTGGAGAGTCTGCGCATTCGCTCGG 1140
QY 1141 AGAGCGCAG 1182
DB 1141 AGAGCGCAG 1182

RESULT 3

SME591792

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

```

/evidence=not_experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC47405.1"
/db_xref="GI:15075851"
/db_xref="SPTREMBL:O92M30"
/translation="MGEAVEMEFVEFPSPASQRIPEVFNSEHSGRHYPOFFLOSRLD
PHSTRSEDDHVDLEFSATFLGAPLLRAHPRAFLDVNREPELDPMDGALPPHA
NISMRYAGIGVPLVRAENMEIYRGRFVEQALERIEIYKPYHATLKRLARTIV
EFGMSILIDCHSMGNVHLPGSGHRRPFIIGDRGTSAAEISRAVELLEOGLYAAV
RNKPYAGGFTEHGRPTGRLHAIQIENGLYVDENLILKKPFALDELDAIFIA
LARVEDGAYPLPAE"
1305..1667
/gene="SMC04044"
1305..1667
/gene="SMC04044"
/function="miscellaneous: not classified regulator"
/notes="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE 2-COMPONENT RECEIVER DOMAIN PROTEIN"
/protein_id="CAC47406.1"
/db_xref="GI:15075852"
/db_xref="SPTREMBL:O92M29"
/translation="MTRKIIAEDNDMRFLYKALEKAGKYIYSTNGASDAIDRLRE
EPFSLITDIDVPEMDGIELARRATEIDPLKWFITGFAVALNPDSKAPKQKVLIS
KPFHLRDVNEVNMMLA"
1952..2026
/gene="tRNA-VAL_GAC"
1952..2026
/gene="tRNA-VAL_GAC"
/product="tRNA-Val"
/notes="codon recognized: GUC; predicted by tRNAscan-SE"
/evidence=not_experimental
2459..4297
/gene="tLYD2 OR SMC04045"
2459..4297
/gene="tLYD2 OR SMC04045"
/EC_number="4.2.1.9"
/function="small molecule metabolism; amino acid
biosynthesis; isoleucine/valine"
/notes="Product confidence : probable
Gene name confidence : probable
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE DIHYDROXY-ACID DEHYDRATASE PROTEIN"
/protein_id="CAC47407.1"
/db_xref="GI:15075853"
/db_xref="SPTREMBL:O92M28"
/translation="MPAYRSRTTGHGNNMAGARLMATGKMSDECKPIIAVNSFT
QFVGHVHKDGLGVAREIEAGVAKKEFTIAVDGIMAGDMGLYSLPSEIITD
SVEVYNNAHCADAMVCISSNCKITPGMLAALRLNIPAFVSSGPMAGKYVILGKTH
ALDLVMAVAADKVSDEVDQIITERACPTGSCSGMPTANSNCTETELGLSLPN
GSTLATVHADKRLFEAGHLVDLARRYEQEDERYLPRIATKOAFENMALDIAGC
GSTNVLVLIILAAVEGIDFTMDIDRLSRKVPCLSVAPAKADVHEDVHRAADGMS
ILGELDKGLINDCPTVHAETLGDALDRMDITRTSDTVRKFFRAAGGIPPOVAS
QEARMELODRENGVIRSVHEPESKDGGLAVLKGNTALDGCIVKTAGVDESLIKPESG
PARVESODAAVVGIIANETKADGVYIRKRGKGPAGMCEMLYPTSYLSKSLGKAC
ALITDGRSGTSLSTIGHVSPRANGTIGLVREGMDIDIDIPRTISLRVDEAELA
ARTEQDAQMKRVEQRKRRTTALKAFAFASADRGAVRDLGDR"
4343..4400
/notes="Sm-5 OR SMC04615
REPEAT SM-5
Predicted by Homology"
repeat_region

```

```

/evidence=not_experimental
/complement(4470..4874)
/gene="SMC04046"
/complement(4470..4874)
/gene="SMC04046"
/function="miscellaneous: hypothetical/global homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC47408.1"
/db_xref="GI:15075854"
/db_xref="SPTREMBL:O92M27"
/translation="MKTKHGSCHGRIREVVIDLEAGTSRCNCISGSLRWGASV
KPDFEFLMCEAGIGDYQFGTMSGHRFCYACGVTPGHGYVEIGAFVINVACID
GVPAEFPAALPIQYMDGLHNNMNAPETRHM"
4979..5024
/notes="REP (repetitive extragenic palindromic) element;
Predicted by Homology"
/evidence=not_experimental
5027..5084
/notes="Sm-4 OR SMC04584
REPEAT SM-4
Predicted by Homology"
/evidence=not_experimental
5094..5542
/notes="Sm-2 OR SMC04873
REPEAT SM-2
Predicted by Homology"
/evidence=not_experimental
/complement(5586..6029)
/gene="azuz2 OR SMC04047"
/complement(5586..6029)
/gene="azuz2 OR SMC04047"
/function="small molecule metabolism; energy transfer;
electron transport"
/notes="Product confidence : probable
Gene name confidence : putative
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PSEUDOAZURIN (BLUE COPPER PROTEIN)"
/protein_id="CAC47409.1"
/db_xref="GI:15075855"
/db_xref="SPTREMBL:O92M26"
/translation="MKTKMILAMLCATSAGOANAEERYEMLNKAADGRVMAFEPAY
IRAQPGTYFVAKDKKGNLSALMKKGARBEARWTKKINETITVLSKPVYMYQCAP
HVGKMTGALIVESEPALNEAVKIGIKTPGSKAAAEKIFATESGG"
/complement(6076..6417)
/gene="SMC04048"
/complement(6076..6417)
/gene="SMC04048"
/function="small molecule metabolism; energy transfer;
electron transport"
/notes="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE CYTOCHROME C PROTEIN"
/protein_id="CAC47410.1"
/db_xref="GI:15075856"
/db_xref="SPTREMBL:O92M25"
/translation="MTHASSRIIAAVLTCVTIGLSPPMASAEEDKIALGREIFLSE"

```


			(DXP REDUCTOTISOMERASE)"
		/codon_start=1	
		/transl_table=1	
		/product="AGR_C_4736p"	
		/protein_id="AK88334.1"	
		/db_xref="GI:15157812"	
		/translation="MTNASEMPKRLTITLSTGSLGTNTLDVPHOLGGROGEFEMALUTG AGNIALIAEOARERGAQLAVYADDDKYEALKSALAGCIIVAGMAGLEAASMDGM VMAIATGPGLAPFTLLTAARRGADIALLANCCUYSAGDVFLRTYKQSGRIILPDSEMS AIFCCLLTGEYKQAVERIVLTRAGGFPTWSDENSVYTAIDIANHPWMSGLVSTIS ASMKNKELEMIEAKYLTDLPDPDQVDVYHPQSIIHSWGVTDSYLAQLSSPMRRAI SYALTPEGRGNLSVERLDPAFKLRDLFEADEDERFPALRIARALLEGIGLOGLAAUNA ERTFFHFVAVGGIGFLDMAELIVETVMRDHGRATMETMDVFSADEEARHALELIAT KERKA"	
gene		complement(7480..8751)	
		/gene="AGR_C_4738"	
CDS		complement(7480..8751)	
		/gene="AGR_C_4738"	
		/note="5-AMINOLEVULINIC ACID SYNTHASE (DELTA-AMINOLEVULINATE SYNTHASE) (DELTA-ALA SYNTHETASE)"	
		/codon_start=1	
		/transl_table=1	
		/product="AGR_C_4738p"	
		/protein_id="AK88335.1"	
		/db_xref="GI:15157813"	
		/translation="MTVPVSEKRTKWSILEHGDMFEAFETTELQSLISEGRYEVFADI EROOGNFPRATRYANGDRKDQVTKESNDYLGMOONKVEIAKKALIDHCAGAAGTR NISCTNHVHLIEOELADLHGKESALIFTSYGVSNNATLGTLOGKIPLGIIFDALNH ASMEGIRGYRCERVIWKHNHLEDIELAKRIADPNAPKILAFESIYSMODDIPIREI CDLADRGMATYLDENVAVGNYGPRGGIAREBSGLDRLLTIESTLYCAKWVGVIYT GSTVAOCFIRSFASGLETTLAPPSLAGIASIOHLKASPFERAKHDVRRLRGILL DARIIPHMDSHIVPVMGVDAACKWISDILLDNHGVOVPINYPVPKTERLRIT PTPLHTADIOLIEUALGHOLMSICALARAVA"	
gene		9017..9319	
		/gene="AGR_C_4740"	
CDS		9017..9319	
		/gene="AGR_C_4740"	
		/note="hypothetical protein"	
		/codon_start=1	
		/transl_table=1	
		/product="AGR_C_4740p"	
		/protein_id="AK88336.1"	
		/db_xref="GI:15157814"	
		/translation="MGRLPSPRRACNVGESFMKKATIFALVGLSIASCOTBEGASTG AVSSAIVGAYTGAVNRCAVAAGAILGVGGALIGNASEPGICYTRDOYGOATYLRCR"	
BASE COUNT	1974 a 2842 c 2765 g 1855 t		
ORIGIN			
Query Match	36.7%:	Score 434.2;	DB 1; Length 9436;
Best local similarity	61.9%:	Pred. No. 2,8e+40;	
Matches 688;	Conservative 0;	Mismatches 423;	Indels 0; Gaps 0;
Oy	4 CGCACGCTGTGCATCTTTGGGGGCCACCGGCTTCATCGGCGAATCCACCCTTGCACCTCGTC 63		
Ddb	6275 CGCAAGCTGACCATATTGGGTTGCGAGGGGCTTCATCGGCACCAATAACTCTGGAATGTCGTG 6334		
Oy	64 ATGCGGAAGGGCGGGCCCGGAGGCGTTCGCGACCGTCGCTGCAGCCGGCGGCGCAATC 123		
Ddb	6335 CGTCAACTGGGGGCGCGAGCGGGTTCAGAGATCATAGCGCGCTGACCGGCGGGGAAAATC 6394		
Oy	124 CGGCGATGCGCCGAANTGGCGCGTGCCTGTGAAGCGGAGAGCTTGGCTTACCCGCGCATGAG 183		
Ddb	6395 GCGCTTTCGGCCGGAACAGCGCCGGGAATTTTGGCGCGCAGCTCGGGGTACTCGGGGAGCAT 6454		
Oy	184 GACTGCTGCCCCGGCTGCGGAGGCGCTGCGCCGGGACGGGACCCGAGGTCGCGGGCGGG 243		
Ddb	6455 GATAAATATGAAGCGCTGAATTCGGCGCTGGCAGGACACCGGCATCAAAAGTGCAGCGG 6514		
Oy	244 GCGCAGGCGCATTCGCCGAGGCGCGCGACCGCGCGCGCGCATGACCATGATGCGCATGTG 303		
Ddb	6515 ATGGCCGGCCTTGAAGAAGGCAAGCTTCGATGGATGATGCGCGCTGGTGTATGCGCCGCAATTGCC 6574		
Oy	304 GGCGCGCGGGGCGCTGTGTCGCCGGAATGCGGGCGGCTGGAAGCAAGCGCCGCACCTTGGCGCTC 363		

CDs

2849. 3151
/gene="Atu2614"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA143595.1"
/db_xref="GI:17741113"
/translation="MGRLPSSRAHNCNGSEFPMKRAIIFALVGLSLASTQCHENKASIG
AVSGVIGAVGATGVNCRGAACVATIGVGALIGNASEPGCYRDOYGOQRYTACR"
3399. 5252
/gene="Atu2615"
3399. 5252
/gene="Atu2615"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA143596.1"
/db_xref="GI:17741114"
/translation="MTFALYPAFARDAFAFLFGMRLLGSESEPEVEVLPNKYAVTLD
ADAKSKLSLENSLLADKRPASDGLILKARDDRRLIAALYENKAYGVIVN
VYAKKNADLPNPNVPFHSTVPVWVITVTPGPKPTGNGLEDDYGRNIDEGILN
GVDSAKIATIRGNKLTIDDLKAREGRPLAKLRREVAHNHATNPDTTAAAGGVATP
GAVTYGKTYVDGDFTRTSRLNGGEPTSPKLRKAADRLQVLQFSSLTIGKATGTA
RGUPLPLTEVSEGHRYFGVAQYSTEIGLQGWGRNLFGQASLRTEGVSRIR
AEASVGEVDVSAGITPTPKPGMFPNRTTFKTSIAKTENPDTYRAKTLTGAGVAYEL
NPTDPAAGLEVOYADTDAFGKNEYLTTSIPLEFRPTRODKLNPTGPFASIAAKR
SEALNGLFESEFEGSITGYGIGLAKEDLTMAGKLSGCVLWAGSSDLIDIPTRPFAG
GGGSVKGYSYQDTSIPYNNAGDAGTGRSVYCSVPRKRYVTPTICLVPPFDAGVSDGV
TPFSDIRAGAGIGLRATPPGRLKLDVAMPLEKTDGNGNITAGIGQSF"
5814. 59458
/gene="Atu2616"
5814. 59458
/gene="Atu2616"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA143597.1"
/db_xref="GI:17741115"
/translation="WRLSLVTAHRRNADAEVNADVAFLPNENVLKAEKMEPEGGL
ATLLSPETPAVALDLNDEGLPSWVTGLRGNVAGNAVYNTGRLHGGDTRRIETIA
GGGGDLPLPAFRQLPAGETKLLANALTYPOGHEIETSGLEAGLLTNGSTVDPN
GONDAANLITAGVDPDRMPPLNGEINALLNGDLSLKGADAAHNNTYVLSLAL
PGRDIVGKTRAESADLNTNRSGTIAVIVSSTPPIDIDILVRAFLTRKPLSTLS
TSAIFGNKATLESASVCGTNGSFDLANLTJSSVQFALPALTLPALAKEPTTTA
LOGDIDLTGGRTSIEHLVKSGTIEAGNVSLENDALISADLKGKPALEKPTTQAG
IADRFILSANGVAADPNVTLSSEBAILAGKRLKALVNASGAKDPAPAKPFASSDL
LADKAIENANVAVOTENGTAPELHVAVGNRIWLGKLOFSSQPLPTNLSNPEFALL
LAALAQAQADGIGIADGLISNANGKIATIKANGINOCGTTTISKLAISIDIDKFA
LAINKISAVNAGTATAISGLNNTIGHSGTTPFDVNGRDNAPLVKSGAIDSGP
MTVRDTSAPKPGIPIVLEKPTSTAIQNGRTARIADULTITGDSREYVNGAGSTLDI
KADISILASLANAFACLDLAGSISGVSFKAGASDSVDYVNLNMANAEAGQRAAG
LAALIKRANGFRAGGNLQIDIDTVGQGGSLSGGSGSLGIANRPLSAMAFSGRLPFSVA
AAQTAQGLIDVGTALDIVNISGTASAPVYVGSITJTGTTGRTLTDRNLITINGLATAT
FDRDAVYLSRLKGLVGGGTTSGTSVYIAGSGEPADISTTLDRAGYNGDTLVTTVV
SGTLTKLKPILNSPALGNTLIDRSALITPKLASLETIDYKRRKNAPRYRAQAKL
CADQGGSSSSTINLDIOVNAAPSGITVFRGKIDELGNTLITKGTAAVPIVSGFEKRR
RGREILITRRLDFTTGNTTFFGGGLVAPVLDKMAKDVSTVTSVSGONADNPFAFESS
APALPODEVMAOLJFGQSSKLSIALQILRLADAAQIAGGSTSLPFLKLSNGLVDLID
DISTSECOARSAAGKYINERTYIELQSGSGSAAKIINLDVGNVGLRBDAGSNGSG
AAGIFYEKY"

gene

9612. 10103
/gene="cheW"
9612. 10103
/note="Atu2617"
9612. 10103
/gene="cheW"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

CDs

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 688; Conservative 0; Mismatches 423; Indels 0; Gaps 0;	61.9%; Pred. No. 2,76-40;	434.2;	DB 1;	10301;
<pre> /codon_start=1 /translation="chemotaxis protein" /protein_id="P4AL3598.1" /db_xref="GI:17741116" /translation="MGRILMAMINSTNGETTEILIAFLRHDOEFCVKTITIREINQM APSPPIPHADVDVIGVNNLMSVPIPIIDLHKIGMSTVANEISAIYVAEVHNVIGM LVDSVDILITIPANQOVPEPISAFSKSEGLIANEHGMICFLNLAKMKFCEAD LAA" </pre>				
BASE COUNT	2178 a	3136 c	2967 g	2020 t
ORIGIN				
QY	4	CGCACCCTGTCGATCTTTGGGGCCACC	CGGCTCCATCGGCGCAATCCACTT	TGGACCTGCTC 63
DB	107	CGCAAGCTGACCATATTTGGTTGCGAGGGGTTCCATCGGACACCAATACCTGTAATGTCGTG 166		
OY	64	ATCGGAAGAGGGGGGCGCCGAGAGCGCTTCCGCGACCGCTGCTGACGCGCGCGCGCAACATC 123		
DB	167	CGTCAACTGGGGTGGCGGAGCGGGTTCCAGATCATAGGGCGGTGACGCGGGGGGGAACATC 226		
OY	124	CGGCGCATGCGCCGAATGGCGTGCCTGAAAGCGGAGCTTGGCGGTACCGCGCATAG 183		
DB	227	GGCGTTTCGGCGGAACAGCGCGGGAAATTTGGCGGGCAGCTCGCGGTACGTCGGAGCAT 286		
OY	184	GACTCTCTGCCCGCGCTCGCGAGCGCTGCGCGCGGAGCGGACCCGAGAGTGGCGGGCGG 243		
DB	287	GATTAATTTGAAGCGCTTAATTCGGCGCTTGGCAGGCGACCGGGCATTAAGTGGCGGACGC 346		
OY	244	GCGCAGCCATCGCCGAGCGCGCGGACCGCGCGCGCGCACTGAGACCATGTCGGCATGCTG 303		
DB	347	ATGGCGCGCTTTGAAGAGCGACCTCGATGATGATCGCGCGCTGAGTGAATGCCCATTGCC 406		
OY	304	GGCGCGGGGGGCTGTGCTCCGGAATGGGGGGGCTGAAGCAGGGCGGACGCTGGCGCTC 363		
DB	407	GGCACCCTGGCGCTTGGCTG 466		
OY	364	GCCCAACGAAGAAAGCCTCGTGAAGCGGACGAGGCAACTCTGATGCGGAGCGCGCCAGAGAAC 423		
DB	467	GCCACAGAGAAATGCTGTTGGCGGGCGCATGTTTTCGCGCATCTCAACAGAGGC 526		
OY	424	GGCGCGAGATCTCTGGCGGTGAGACGAGCACTCCGGCGCTTTCAGGCGCGCGCGGCG 483		
DB	527	GGCGCGCGCTGATCCCGGTGACGACGCAAGCGCAACTTTCAGTGCCTGACGCGGT 586		
OY	484	GAGGACAGCGGCTGTGTGAGCGGCTCATCATCAAGGCGTCCGGCGGGCGCTTCCGCGAC 543		
DB	587	GAATATCAAAAGCGGCTGGAACGATCTGTGTGACGGCTTCGGGGGGCGCTTCCGCGAC 646		
OY	544	TGGAGCTTCGAGCGCATCCGCGCTGCGACCGTGGCGGAGGGCGAGCGCCATCCCACTGG 603		
DB	647	TGTGTGCGGGAGAGATCTCCACGCTCCGCGGATATGCTCCCGGCCATCCCACTGG 706		
OY	604	TCCATGGGCGAGCGGATCTTCATGACAGCGCGCTTCGATGTTTCAACAAGGCGCTGAGCTG 663		
DB	707	TGCGATGGGGCTGAGGTTCTCGATCGCGCGAGTGCCTTCATGTTTCAACMAAGGCGTGGAAATG 766		
OY	664	ATCGAGAGCGCGAATTTCTTGGCTTTCGAGCGCGGACCGGATCGAGGCGGTCTCCATCG 723		
DB	767	ATCGCGGGAATATATCTTTCGATCTCAGGCTGATCAGTGGAGCGTGAATGTGCTCATTCA 826		
OY	724	CAATTCATCTGTCATGCGATGTTGGGCTTCTGCGACGGGGGCGCTGATAGCCCATCTGGCG 783		
DB	827	CAATGATCATTCATCTTCATGATGGTGGCTAACCTAGACGGCTCTTATATCGCGCACCTGGC 886		
OY	784	CCCGCGCATGCGGCAAGCATCGGATTTCCGCTGGAAGTGGCGGGCGGGGAGAGTG 843		
DB	887	TCTCCCGCATGCGCACCGCCATCTCTCATATCCCTCACATCTTCTGAGCGGGCGAATCTA 946		
OY	844	CCCGTGGCGGATTCGACCTTGCACAGATTTGCGAGCGCTCACCTTTCAGAAAGCGTACGAG 903		

Db 947 AGCGTGGAGCGCTGATTTCCGCAAAATGCGACGCTCATTTCCAGACCGGAGAA 1006
QY 904 GAACGCTTCCGGCCCTGAGGCTTGGCGGAGAGATGCGGGCGGCGCTGCGGGC 963
Db 1007 GCGCGCTTCCGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1066
QY 964 GCGCGCTTCCGGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1023
Db 1067 GCGCGCTTCCGGCGGCGGCGGCTGCGGCTTTCACGCTTTCGCTGCTGCTGCT 1126
QY 1024 TTTCTGACATGCGGCGGCTGCTGCGGAGACGCTGCGGCGGCTTTCGACCGACCCCTG 1083
Db 1127 TTTCTGACATGCGGCGGCTTTCGAAACGCTGATGACGCTGATGACGCGCGCAGC 1186
QY 1084 TTTGGAAGATGCGGCGGCTTTCGAGGAG 1114
Db 1187 GCGGAGACGATGAGCAGCTTTTTCGCGCAG 1217

RESULT 6
AE005865/ 12977 bp DNA linear BCT 12-JUN-2002
LOCUS AE005865 AE005673
DEFINITION Caulobacter crescentus CB15 section 191 of 359 of the complete genome.
ACCESSION AE005865.1 GI:13423361
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 12977)
Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Uterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and Fraser, C.M.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Complete genome sequence of *Caulobacter crescentus* Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
2 (bases 1 to 12977)
11259647
Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Laub, M.T., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Uterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Shapiro, L., Venter, J.C. and Fraser, C.M.

Institute for Genomic Research, 9712
D 20850, USA

04:26 2003

protein family: HM

/translation="MTINRDKPLKYLVAEASDGLAGLAKLARLGGVTEVGI
GAKMAEQVQSPEDTIAQSLIIGLISLAKYAPRAARLKDVTALAREKPDVALIDS
WGENFALRALRDLPTLPLVYVPOVWVAEGRQALAKAVDLISLTPDKAYFD
AGLQNVFVNSALAKRDPHADPGRRLAIGAAISOILLVPSRSPSEIEVMAFE
DAVRLKVERPDLHYVPAATYVAEYKARVAPRAHYIEDEGKDAFLAVAPAE
ACSVITTELALAGRPVAVGTGATTAIVRLKMPRTITLNLNLAAGAAVAPLLOH
ACEGELAREVALRLDDPDLRQRTAEYALDRMGKMDPESEAAEALDLFLSARG
ALGRAG"
/complement(1267..2109)
/gene="CC1910"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK23885.1"
/db_xref="GI:13423363"
/translation="MRKGLIAGGALPYELASHCEAAGRAVNRSLRSPADSDRY
PADVIGEFKIFKALRAEGDVCYFAGNVRPDPESALRMPARGIKYPLILVARK
GPDALLRLVDEPEKRGFRIGAHVEMGEMTLPRGLGVSPAPERMADIDKLVAR
EIGRLDIGGAVYCSGLVAVAEQSTDMALRVADLPRAIRGARRRIGVLAAPKAP
IDETRVLDPTIGVATLIRARAGLAGIVGAGRLLVVREAVIAAADLGLVLDVDP
QERP"
/complement(2119..2910)
/gene="CC1911"
/complement(2119..2910)
/gene="CC1911"
/note="identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="acyl-(acyl-carrier-protein)--UDP-N-
acetylglucosamine O-acyltransferase"
/protein_id="AAK23886.1"
/db_xref="GI:13423364"
/translation="MSINHPALIAPEAKLAPVEVGPESIVGPDVTLAAGVRLSHV
VQATVIGGCVVSHFANLGGPPIHAGKERTELIGRNIRIIEHVMTGTASGK
VTTIGSDGLKVMGSHVAHDCVDFVLAKKATIGGVAIGDVGGLAAHOFERI
GRSFGLAATVATKDYIPYGSWGNHAEGLNVLKRGPPREINLRLALRME
ADEGTFQERLDVATIEHAGNAEMELVDIRIDANRPLDLPERV"
/complement(2907..3386)
/gene="CC1912"
/complement(2907..3386)
/gene="CC1912"
/note="identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="(3R)-hydroxymyristoyl-(acyl-carrier-protein)
dehydratase"
/protein_id="AAK23887.1"
/db_xref="GI:13423365"
/translation="MGDNAEQAVNDIDIAELIARIPHRYPPLVNDARDYNPHOSIV
GIKCVTINPEFPOGHPGNPMPGVILITLALOTGAVLMSKSLVDUTKEFTFMSK
NAERNPVRPGDVRMEVEVLRARSSIFEFKGVAKGKVAEAEPAAVVETGPRA"
/complement(3389..4408)
/gene="CC1913"
/complement(3389..4408)
/gene="CC1913"
/note="identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="UDP-3-O-3-hydroxymyristoyl glucosamine
N-acyltransferase"
/protein_id="AAK23888.1"
/db_xref="GI:13423366"
/translation="MPDPREFDSLGPALLISELAQAATLADALGERVITTAAPLDA
SDAQATTFESDAKRDAASTRAGCFVPRHQGFPPICALVYGRPAAMAAANR
LHAPRRHAGAPSLHPDALEDEVALAPVITIGOGASIGRTIGGVVIGGVVIGR
YCRIGANAVIGFAMLDNVAISAGVIGAGFGAALGPRGVLDLPOLGAVVIODNTL
GANSVDRGAFGPTTIGENTKIDNLVHAVHANYRIGENCYLAAYTVGSGSTVVDGVA
GKAGVADHLNIGSGASIGAAASVFEVDVDEGTWICFPARPLKRLRLRETNALSRAAG
RGTG"
/complement(4462..5100)

gene

...identified by Glimmer2; putative"

...start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK23889.1"
/db_xref="GI:13423367"
/translation="MSKFLPAASVVALAASSASNOTAPAPAPAVYTHGPAALPG
VCIFSSPRAVSSLVGKAVDARLKTIIQVNAELTGERTALDNEAKALDAKKTIIAD
ALFOQATLOAKANAMOROKLOKEVEAEFOKALSHVOELNPTIOOYVOAKCSVL
LDREAVLAMPANDIDAVYALDARKITITPDERLDQVPGAAALQPTNK"
complement(5103..7412)
/gene="CC1915"
complement(5103..7412)
/note="identified by match to PFM protein family HMM
PF0103"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAK23890.1"
/db_xref="GI:13423368"
/translation="MALLGSTALVAPQAQAOTGVVORITIVGNERIEGTVLS
YLPIGDTVDQSORDLAKTLARPTDLFAVKTLEMIGDGLVYKVENPITNOYFEN
SSIAKEDKLAEVOIRPRGIFTRAKVQADVDRIELRRSGRISITVPKVELLPQKEV
DLVEINEGAKSVLGINFLGNAEYSDNDRIYIVKESWYILTSNDYDDRIEY
DRQLRKHYNRGYDFEYVSSVAELAPDNGFAVYITLEGRKXGKLTIVETELK
LDENLAQILVPTGOLYEDERIEQATDALTPAAGAAGFNVYRPRYVNRKTKVD
VFOVREGPRVYRIDIVNTRTLDVLRLELVEAGDVAVNLYDRSKNNMRIGF
FKEVEIDAGSAPDRSLRKYVEOPTGELSPSAGYSDTKLVLDVQITERNPRGQ
QNRARASVGSLSQOIDFCSEPRFLGRNLYAVGNLTFRIDISERAYDTKSVGGV
RFGPLTNDSSMSLRVYRQDEVSVADSLCASVSOILCLOGAVITSLIGGLRID
KRNDIPNTRGWEADLNQDLAVGQDVYKLTADAGWYGFKDIYFASGSGYIE
GMCQGVRRINDRFYRGTSFSGFEIAGIDRDISSFSNMGAKLYAISTFELVPTFL
PEQYGIKALFSDVGTAGLDDVDYDRSPVFPNPKNDNLGLASAGISIDWMSPMK
IRPDISILSKKEYDRETFRSTSTRFQ"
complement(7484..8680)
/gene="CC1916"
complement(7484..8680)
/note="identified by match to TIGR protein family HMM
TIGR0054"
/codon_start=1
/transl_table=11
/product="membrane-associated zinc metalloprotease,
putative"
/protein_id="AAK23891.1"
/db_xref="GI:13423369"
/translation="MIGFLIMLVSLFVLSVVTVEHLHYAARACVAIERFSIGF
GAPLISWRDRKRGVENCVASIPLGIVRFADENAAVPPDNDLDAKNEIRREGDA
VNRYPFRKPYMORAFIAVAGPANFLIALIVFVIVISFAQCTSTVTPKVELLPQKEV
AAFGKQDVLKADNRQIDSFQDIQGVIALRANPIDEAVERGVYVHLTAETRLER
ONEISGRVKGVELGLRSAPGREGFSSLSAIVDAVEWDMKTTAFYATRLMOL
PADQISGIIIGHTAGAVTNGVYEQAPNKAALAGIISYFMIASLSISIGMNLIP
IYVLDGHLVMTAYEAVAKRPLAEFQAGFRAGLALILGEMFLPAAAMNDLNRVDERF
IGGLFT"
complement(8703..9902)
/gene="CC1917"
complement(8703..9902)
/note="identified by match to TIGR protein family HMM
TIGR00243"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AAK23892.1"
/db_xref="GI:13423370"
/translation="WGLTSPKRVVVLGSGTSLSLFEESGAPVOIALTLAQRN
VERLIOARWKPMSLAVIEDESLDLRAGLGTGVEAAGADAVDAAMGADWYMS
AIVGAGLAPVAAARTGAVIALANKESIVCAAGPALATATAKAGSGVTPDSHSAIF
OYVQSECAHRVSRLLITASGGPRTWDKAMAMARATEQALAHNMMSGAKISVDSATM

Query Match 29.8%; Score 352.6; DB 1; Length 12977;
Best Local Similarity 59.0%; Pred. No. 3.4e-31;
Matches 626; Conservative 0; Mismatches 429; Indels 6; Gaps 1;

MNKGLEMIENASYLPATPEDEVYVIRHPOSYIHSVLYVDSGTIAQLGPPMRAPIACA
FAMPDRLPMPARLUDLAATGQLTFESDEPRFAIGALRALGLGAPRANMAAEV
AVAFEDLRIGFLDIAGAVAGLEIRNMNSLDLSVASDAETMLIDGSAIRIAAEV

QY 4 CGGACCTGTCGATCTTTGGGGCCACGGGCTCATCGCGGAATCACTTCGACCTGTC 63
DB 9881 CGCAAGGTGTGGTGTAGGCTCAGCCGAGATCATCGGCTTTCGACCTGACCTTTC 9832
QY 64 ATGGGGAAGGGGGGGGGGGAGGCGTTCGCCAGCTCTCTGAGCGGGGGGGAACATC 123
DB 9821 GAAGAGTCGGGTGGC-----CGTTTCAGATCTTGGCGCTCACCGCGCGCAATGTC 9768
QY 124 CGGCGACATGCGCGCAATATGCGCTGCGCTGAAAGCGGAGCTTGGCGTACCGCGCATAG 183
DB 9767 GAGCGCTGATCGAGCAAGCGCGGGCGCTGGAAGGCTCTCCAGCGGCTATCGAGACGGA 9708
QY 184 GACTGCTGCGCGCTGCGGAGGCGCTGGCGGAGCGGCGACCGGACGATCGCGGCGG 243
DB 9707 AGCGGCTTGAATGATCTTCGCGCGGCTGCGCGGCTGCGGCGGCTGAGAGCGGCGGCGG 9648
QY 244 GCGGAGCGCATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 303
DB 9647 GCGGAGCGGCTTGGCGATGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9588
QY 304 GCGCGCGCGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 363
DB 9587 GCGCGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9528
QY 364 GCGCAAGAGAAAGCTCTGTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 423
DB 9527 GCGCAAG 9468
QY 424 GGGCGCGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
DB 9467 GGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9408
QY 484 GAGGACAGCGGCGCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 543
DB 9407 GAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9348
QY 544 TGAAGCTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603
DB 9347 TGGGACAAAGGCGGCTATAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9288
QY 604 TCCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
DB 9287 TCGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9228
QY 664 ATGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
DB 9227 ATGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9168
QY 724 CAATCGATCGTCATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 783
DB 9167 CAATCGATCGTCATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 9108
QY 784 CCGCGCGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843
DB 9107 CCGCGCGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9048
QY 844 CCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 903
DB 9047 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8988
QY 904 GAACGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 963
DB 8987 GAGCGGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8928
QY 964 GCGGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1023


```

Db 8927 GCGCGATGAACCGCGCAACGAGGTGGCGTGGCTGCTTCTTGACCGCCGATCGGG 8868
QY 1024 TTCTTGACATGCGCGCGGTGTCGAGGAGACCGCTCGCGG 1064
Db 8867 TTTTCGATATTGCGCGCGGTGGCGGACCGCTCGGAGCG 8827

RESULT 7
AE012791/c
LOCUS Chlorobium tepidum TLS section 12 of 194 of the complete genome. BCT 01-JUL-2002
ACCESSION AE012791 AB006470
VERSION AE012791.1 GI:21646016
KEYWORDS
SOURCE Chlorobium tepidum TLS.
ORGANISM Chlorobium tepidum TLS.
Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae; Chlorobium.
REFERENCE
1 (bases 1 to 11150)
Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M.,
Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H.,
Hickey, E.K., Peterson, J.D., Durkin, A.S., Kolonay, J.L., Yang, F.,
Holt, I., Umayam, L.A., Mason, T., Brenner, M., Shea, T.P., Parksey, D.,
Nierman, W.C., Feldblyum, T.V., Hansen, C.L., Craven, M.B., Radune, D.,
Vamathevan, J., Khouri, H., White, O., Gruber, T.M., Ketchum, K.A.,
Venter, J.C., Tettelin, H., Bryant, D.A. and Fraser, C.M.
The complete genome sequence of Chlorobium tepidum TLS, a
photosynthetic, anaerobic, green-sulfur bacterium
Proc. Natl. Acad. Sci. U.S.A. 99 (14), 9509-9514 (2002)
12093901
JOURNAL PUBLISHED
2 (bases 1 to 11150)
Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M.,
Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H.,
Hickey, E.K., Peterson, J.D., Durkin, A.S., Kolonay, J.L., Yang, F.,
Holt, I., Umayam, L.A., Mason, T., Brenner, M., Shea, T.P., Parksey, D.,
Nierman, W.C., Feldblyum, T.V., Hansen, C.L., Craven, M.B., Radune, D.,
Vamathevan, J., Khouri, H., White, O., Gruber, T.M., Ketchum, K.A.,
Venter, J.C., Tettelin, H., Bryant, D.A. and Fraser, C.M.
Direct Submission
Submitted (30-APR-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
unpublished
FEATURES
source Location/Qualifiers
1. .11150
/organism="Chlorobium tepidum TLS"
/strain="TLS"
/db_xref="taxon:194439"
322. .1779
/gene="CT0117"
322. .1779
/gene="CT0117"
/codon_start=1
/transl_table=11
/product="sulfide-quinone reductase, putative"
/protein_id="AA071365.1"
/db_xref="GI:21646017"
/translation="MAKVVVLGAVSGHSCAGSLKKLKGQHEVVVSPNSYYQWIPS
NTWVGVMHTIDVREFKLAKVYDRGIDYKQAKVSIHPEGDANISKGYVTLEYTDEE
HAGYGTVDYDLYNATGPKLNFATEGLGPKNSLSVCTYSHAAHAELOKSIKEM
KNGKQRFELIGHAMATCGAAFEYILNVAHEISRRGLSHMAELTWISNEYLDGFG
MGAFTRKGYITPTKVFTESLIAEYIKWIRAGYKVEPGVAHYETLDGMLSQEF
DFAMLIPSGVGLTAFDKSGNDITDKMFLPNKFMKVDADYTAKPFGEGANDWPTIY
QTPMYSNIYAAGTAFAPPHSISKPTSVNGROIPTTPPTGMPGSGVIGKIIALNISEQ
IKGNKHEHHKSMARMGAACIVSAGFSFGDLGASMTVEPVPDWEKYPEWGRDWTY
SVEVGLAGHLKFLMLHLFFHAKGYPFWLIPE"
1830. .1988
/gene="CT0118"
1830. .1988
/gene="CT0118"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

```

```

/product="hypothetical protein"
/protein_id="AA071366.1"
/db_xref="GI:21646018"
/translation="MGNFKKAYDEAYPPVDKATLFWRKFLPWOLFPRFFILNIKMI
RIVVGHS"
2090. .2188
/gene="CT0119"
2090. .2188
/gene="CT0119"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA071367.1"
/db_xref="GI:21646019"
/translation="MLLYSVOESHFVIOQLRLAMLVPGVSNKQE"
complement(2352. .2870)
/gene="CT0120"
complement(2352. .2870)
/gene="CT0120"
/note="identified by match to PFAM protein family HMM
PF00881"
/codon_start=1
/transl_table=11
/product="nitroreductase family protein"
/protein_id="AA071368.1"
/db_xref="GI:21646020"
/translation="MTNLHDLAESRKNVSYDTSQPTPOILERVLDANLAPSAKNL
QPTWELLVSSPEMLAKIRSCYKRDMPQQAHPHLLVTDGRDANVWRSPDGMNSIETDLG
IAMDQLIAAHAEGLATCWISAFDPAMLREALDLKPFSEVFAITPGYASADAQTRPK
SRKPLNEVVRYL"
complement(2867. .3898)
/gene="CT0121"
complement(2867. .3898)
/gene="CT0121"
/note="identified by match to PFAM protein family HMM
PF02080"
/codon_start=1
/transl_table=11
/product="potassium channel protein, putative"
/protein_id="AA071369.1"
/db_xref="GI:21646021"
/translation="MSKDAQQISALRRFSVSIISVALLVISTGVGMYLENLSLLDAL
YMTVITATVGFSEVRPDDVGKIFTMVLIVGTTGTFITLTSVAVFFVSGEWKEHWE
QQRNRLKLNDFITICGYGRGLSVAEBELAKAVFVVIDNMDIVNLRARDEGLA
IKGNAADEVLADAGRAKGLIAAGNDAENFIVLTARNLKNPLYIVARADCSE
SKLRAGAEEKVVMYLSACKRMASLLTEPELEFYLDELNNANLNLRIAQYLVGDNSP
LYGKSFQEVLDLYNNHRLNVVGYKLPDGLHTTPRAELIKNCTIIVIGKSDDEMLC
KLAQGETQR"
complement(3935. .4549)
/gene="CT0122"
complement(3935. .4549)
/gene="CT0122"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA071370.1"
/db_xref="GI:21646022"
/translation="MLEMKTGSKQWFFRNFDDTVDYLSDHTRLSYNPECHKVE
PLDRDAYRWHRFVTDPPQNNPDVIFNIQOETELVDLPDEVASMDPEEMSEMIQOF
TVGRKITWTPLAQDKFTTPEKLYFQGTADMLIVPQQEQTRVDLWNVNVAFLLY
PAFRIVPEKVRVMTVSTGMSLIMQTATNMHFKISKEFGKIRKL"
complement(4707. .5780)
/gene="prfa"
/note="CT0123"
complement(4707. .5780)
/gene="prfa"
/note="identified by match to PFAM protein family HMM
PF00472"
/codon_start=1
/transl_table=11
/product="peptide chain release factor 1"

```

```
/protein_id="AAW71371.1"
/db_xref="GI:21646023"
/translation="MEDKLOSIRKDKFQTIQEQSLDEPVVADQNRRFKLNKEYSSSLKEI
VRAYDSMTKKOLDRAHNGKNENDPEMRALVEEAGEQLERLQKLEQQLKLLLPK
DEADSRNAIEIRAGTGDEAGLEAFADLRMYOYAEROGWQSCOTLVEQSGVPSGLK
EVSLEYGNHYGILKPFSGVHRVQRPETEGRIHTSAASAVAVLPEAEVDVEIRK
EDLLIDTFRSGGQGNVKNKVTAVRITHVPSGIYVACQBEERSQLQNRERAMKLRSK
LYDLQAEQKSRADLRSMVTGDRSAKIRTYNFPQSRVTDHRIGFTSHAUPQIMQG
ELDPLEALRMHQDAERLQAE"
/complement(5802..7163)
/gene="Ct0124"
/complement(5802..7163)
/gene="Ct0124"
/notes="identified by match to PFAM protein family HMM
PF00595"
/codon_start=1
/transl_table=11
/product="membrane-associated zinc metalloprotease,
putative"
/protein_id="AAW71372.1"
/db_xref="GI:21646024"
/translation="MELLNTIFFVIAFILVTAHEFHFIITARMFGMRVDRFFIGFD
FWGIKLWQKIGTGYGICAFPIGGYVKIAGMIDESMDTHVSOEVPQPEFRAPVMO
RLIYLAGGVAMVLAIVFIGTISFGESEPTITPAFIEPKSVFSSMGQSGDHLV
AIMGKHLWEELDERLASGLQTIERNGEELTLTAPKDIISRINGNSIGIRPT
VPVIDQVLPGDPAARAGIMPGLITAINSGPVADWSEVNIISANAGKLLVTWHL
KNSGTPLAALIRKKGOTITTEVTNNSGKIGISLKQIETIERIKLSLQALIASLN
QWKTIVLVQVQFGKIFSGKEDFRKSVGGPIKARIANOSAEQGPISFMYFVAVLSIS
LATINILPALDGGQFVLNAIEGIRPEIPVEVKRIQQVGMVLTLLMLFAYFMINDL
LNP"
/complement(7209..8357)
/gene="dxx"
/notes="Ct0125"
/complement(7209..8357)
/gene="dxx"
/notes="identified by match to PFAM protein family HMM
PF02670"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AAW71373.1"
/db_xref="GI:21646025"
/translation="MKSLSILGSTSIGSLTLDVYRRHFERPESIAALAEHVDVEMLLK
QIDFSPISVDRSESRRLKGLMDHPEILICLEGAEAAVADGADVMSAIVGA
AGLVPTVIAEAGKIDALANKETLVVAGOLVSDLVKKHDKVLLPVDSEHSAIFQSLVG
HRTEDIERILTAGSGGPFRTAEELKNVPEQALKHPQWSMGAKITIDSATLMKGL
EVTEAHLDPMAEKIGVYVHPOSIIHSMVEYIDGCVIADLGVPDMRAPYALAWPE
RCTGTGIGLDITKVAITLTPEEDMEREPALRLAFADKAGOTYPAVLNAANEIATAAF
LOKIGFTDIAGTVDTMQAHEAWTPTILEEYLQADKWARETARQLIG"
/8515..8591
/gene="trna-val-2"
/8515..8591
/gene="trna-val-2"
/product="trna-val"
/complement(8634..8912)
/gene="Ct0126"
/complement(8634..8912)
/gene="Ct0126"
/notes="identified by match to PFAM protein family HMM
PF00708"
/codon_start=1
/transl_table=11
/product="acylphosphatase"
/protein_id="AAW71374.1"
```

```
Query Match 25.8%; Score 305.4; DB 1; Length 11150;
Best Local Similarity 57.2%; Pred. No. 6.8e-26;
Matches 602; Conservative 0; Mismatches 436; Indels 15; Gaps 2;

QY 9 CCTGTGCATCTTTGGGCCACCGGCTCCATCGCGGAATCCACCTTCGACCTCGTCATGCG 68
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8349 CTTATCCATTCTCGGCAGTACCGGCTCCATTGGACTCAGCACACTCGACGCTCGTCAGCG 8290
```

```
QY 69 -----CAAGGGCGGGCGCCGAGGGCTTCGACACCGTCGCTCTGACCGGGCGGCCACACAT 122
Db 8289 TCATCTGAAAGGTTTTCGATTGCGGCCCTTTCGCGAGGGCCACGACCTCGAAATGCTCT 8230
QY 123 CCGCGCACTGGCCGAAATGCGCGTTCGCGTGAAGCGGGAGCTTCCCGTCAACCGGCGATGA 182
Db 8229 CAGCAATCGATGAATTCAGGCGCTCCTGCTCGGTGCGTGACGAGGGGTCCCGGGA 8170
QY 183 GGACTGCTGCCCGCTGCGCGAGGCGCTGCGGGACGCGGACCGGACCGAGGTTCGCGGGCG 242
Db 8169 GCGCTCAAGGGGATGCTCGCGCATCAAGCCGGAATTCCTCTCGCGCTCGAAGGTGC 8110
QY 243 GCGCAGGCCATGCCCGAGGCGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302
Db 8109 AGC-----TGAAGTTGCGCGCTCGATGGAGCAGCATGTAGTTTTCGGCCATCGT 8059
QY 303 GGGCGCGCGGCGCTCGTCCCGGGAATCGGGCGCTGAAGCAGCGGCGGCGGCGGCGGCG 362
Db 8058 GGGTGGCGCAGGACTGGTGCCAAACCGCTCAGGGCTATTGAAGCCGGAAGGACATCGCGCT 7999
QY 363 CGCAACAAGGAAAGCCTCGTGCAGCGGCGGCAACTCTGATGCGGACGCGGCCAGAGAA 422
Db 7998 GCGCAACAAGGAGAGCTCGTCTGCGCGCTAGCTGCTTCGACCTCGTCAAAAGCA 7939
QY 423 CCGCGCCACGATCCTCGCGGTGGACAGCAGCACTCCCGGCTCTTTAGGCGCTGCGGGG 482
Db 7938 CGATGTCAAACTGCTGCGCAGTGGACAGCAGCACTCGCGGCTATCTTCAGTCTGCTG 7879
QY 483 CGAGGACAGCGCTCGCTCGAGCGGCTCATCATCAGCGGCTCGCGGCGGCGGCTTCGCG 542
Db 7878 CCACCGCAGGAGATATCAAGCAGCATCTCTGACCGCATCAGGTGGCCCATTCGGTAA 7819
QY 543 CTGAGGCTCGAGCGCATCGCGGCTGCACCGTGGCGGAGCGGCGGCGGCGGCGGCGGCG 602
Db 7818 AACGCCCGCAGAGGTTGAAAACGCTCGCGCGGAGACGCGCTCAAGCACCGCGAGTG 7759
QY 603 GTCATGGGCGAGGATCTCATFCAGAGCGGCTCGATGTTCAACAGGGCGCTCGAGCT 662
Db 7758 GTCATGGGCGGAGGATCACTACGACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7699
QY 663 GATCGAGCGCGGAATCTTCGCTTCGAGCGGACCGGATCGAGCGGCTCGTCCATCC 722
Db 7698 ANTGAGCGCATGCGCTCTTCGACATGCGCGGAGAGAAATCGGCTCGTGGTGACCC 7639
QY 723 GCAATCCATCGTCCATGCGTGTGGCTTCTTCGCGAGGGGCGCTGATGCCCGCATCTCG 782
Db 7638 GCAGAGCATCATTCACTCGATGTTGAGTACATCGACGCGCTGCTCATCGCCAGCTCG 7579
QY 783 CCGCGCGGACATGGCCACCGCATCGATTCGCGCTGAACCTGCGCGGCTCGCGCGGAGGT 842
Db 7578 CGTGCCCGACATCGCGGCTCCCATCGCTATGCGCTTCGCGCGGAGGCGCTCGGAAAC 7519
QY 843 GCGCGTCCCGGATCGACCTCGCACAGATTGCGAGCGCTCACTTCAGAGCGCTCGACGA 902
Db 7518 CGGATCGGCAAGCTCGATCTCACCAGGTGCGCCACACTCACTTCGAGGAGCGGACAT 7459
QY 903 GGAACGCTTTCCGCGGCTGAGGCTTCGCGAGAGCGTATCGCGCGGCGGCGGCGGCTG 962
Db 7458 GGAGCGCTTCCCGGCTTCGCGCTTCGAGCGCTCAAGCGGCGGCGGCGGCGGCGGCTAT 7399
QY 963 GCGCGCTTCAAGCGGCGGAGGAGATCGCGCTCGATCATTTATCGCGGAGCGGCGGCT 1022
Db 7398 TCGCGTACTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7339
QY 1023 GTTTCGACATGGCGGCGGCTGGTTCGAGGAGAC 1055
Db 7338 CTTTACCGACATCGCGGAGCGGTGCGATAAAC 7306

RESULT 8
AL646064 201050 bp DNA linear BCT 07-DEC-2001
LOCUS Ralstonia solanacearum GMI1000 chromosome, complete sequence;
DEFINITION
```

segment 8/19.
AL646064 AL646052
VERSION AL646064.1 GI:17428340
KEYWORDS
SOURCE Ralstonia solanacearum.
ORGANISM Ralstonia solanacearum
Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
1 (bases 1 to 201050)
Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
Chandler,M., Choisme,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schiex,T.,
Sigulier,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.
Genome sequence of the plant pathogen Ralstonia solanacearum
Unpublished
2 (bases 1 to 201050)
Boucher,C.A.
Direct Submission
Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, LMCM CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biométrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex, Laboratoire de Genétique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html.
Location/Qualifiers
1..201050
/organism="Ralstonia solanacearum"
/strain="GMI1000"
/db_xref="taxon:305"
206..1447
/gene="Rsc1326"
/note="RS02848"
206..1447
/gene="Rsc1326"
/EC_number="2.6.1.1"
/function="small molecule metabolism; amino acid
biosynthesis"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE ASPARTATE AMINOTRANSFERASE PROTEIN"
/protein_id="CAD15028.1"
/db_xref="GI:17428341"
/translation="MKTQKSAKLNVCYDIRGPVLEKAKQMEEGHQIILKINIGLNA
PFGDPAPEIQDMIRNLNPSAGYSDSGIFAAKAVMHYQOQGIKNVLDIYVGN
GASLIATLATNALLDAGDELLLPADYPLWTAMTSLSGTTPVHYLCDSNGMPPDLD
IRAKITPTKGVIVNPNPTGALYSDALLHDIATAREHGLVIFADRYDKVLPDDN
RHTANASLSEDLVTFNLSKSYKSRAGVMVSGDKRPADYIEGLNWLMSRL
CANVPGQWAIQTALGGYQSINDLVAPGGMRRORDLAYELITAIIPGVCVKPKAALY
FPRLDPSPYIPIDDDQTFIRQLLEERVLVQGTGTFNWHSPDHFRIVFLPHEDDLEAI
GRIARFLERYQRHGTGTRAA"
1538..2857
/gene="Rsc1327"
/note="RS02849"
1538..2857
/EC_number="1.1.1.3"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE THREONINE SYNTHASE PROTEIN"
/protein_id="CAD15030.1"
/db_xref="GI:17428343"
/translation="MOYSTRGHAEPQSFRRLLGLLAPDGGLYLPEQYQVSAEELT
RWRLPYADLAFELILRKPTDIPEDDLRALTRTYTPEVCNARNGDNTADITPLHRL
GEAGTSLSLCLSGNPTLAFKDMAMQLGNLFYALDREHAELNIGATSGDTGSAA
EYMRGKRGVFMLSPHRKSAFOTAOQFSLQDNIFENIATIEGVDDAODIVKAVSN
DLEYKARKIGITVNSINARVAVVYFKGYLLAAPKIGDKVSCVPSGNGVNCAG
HIARMGLPIDKLVTATNENDVDFRTGTYRRAAQYHTTSPMSDISKASNER
VFVRYGTWDTHTADGVKVAENLVPGVPMVYLETALPAKFGDTIREALDREPERPAG
YADIESLPQFEVMPADADRIKGYIAQHTGL"
4361..4957
/gene="Rsc1329"
/note="RS02851"
4361..4957
/gene="Rsc1329"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD15031.1"
/db_xref="GI:17428344"
/translation="NAEAFASCLHDIRACRICEGALPHEPRPVVLADRRARVLIVGO
APGRIVHETGIPWDRSGDRLEWLRIGYDAFYADRRITAVVPMGCFCTGTGNGDLPP
RPACAPQWHRPLLATAMPDLKLLLVGGYAQRYILDVPRGMPLTEVVRGPGVRRDAGRV
LFLPLPSPRLLAWFKHNPWFDAQIVPLRLAALAAFY"
5037..6296
/gene="moea2"
/note="Rsc1330; RS02852"
5037..6296
/gene="moea2"

Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE HOMOSERINE DEHYDROGENASE OXIDOREDUCTASE
PROTEIN"
/protein_id="CAD15029.1"
/db_xref="GI:17428342"
/translation="MKPIKIGLLGTVGGSTLKVLRQNEIRRRRAGRGIEVALIAY
RNVERARAIVQETGASVPVDPVAVVDSPEIDIVVELIGGFDLRLVLRRAIRANGKH
VYTKALLAVHGENEIKFAAQDKGVMAFEAAVAGGPIIKALRGLGNRIEWIATGI
INGTNFTLSEMRDKGLDFGLAEARQIRGLYAEADPTFIEGFDAAHRLTLMSAIATG
VPVQFERAYVEGITKLEAVDIRYAEELGRIKLLGLTRRAEAGIELRVHPTLVPATRL
IANVEGAMNAVVDGADVPTLYYKGGAGAPTAASAVVADLVVDVTRLTADPEHRVPH
LAFQPDALSNTPLLPIEVRSAIYLMRVADETGYLADITRILAESGISIDAMLQKES
REGEPQTDIIMLSHVVEKQVNAATAEALPTVLKSTRLRKEELN"
2889..4334
/gene="thrc"
/note="Rsc1328; RS02850"
2889..4334
/gene="thrc"
/EC_number="4.2.99.2"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE THREONINE SYNTHASE PROTEIN"
/protein_id="CAD15030.1"
/db_xref="GI:17428343"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

REFERENCE
AUTHORS

1 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Chambergio,F., Clapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,Y.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
Comparison of the genomes of two Xanthomonas pathogens with
differing host specificities
Nature 417 (6887), 459-463 (2002)
22022145
12024217

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Chambergio,F., Clapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,Y.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
Direct Submission
Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil

TITLE
JOURNALFEATURES
source

Location/Qualifiers
1. .10029
/organism="Xanthomonas campestris pv. campestris str. ATCC
33913"
/strain="ATCC 33913"
/db_xref="ATCC:33913"
/db_xref="taxon:190485"
/note="pathovar: campestris"
complement(72. .2528)
/gene="oma"
/note="XCC1365"
complement(72. .2528)
/gene="oma"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="outer membrane antigen"
/protein_id="AA040663.1"
/db_xref="GI:21112425"
/translation="MKLLHSPRLSSVAGTQOHPQSTPTGCDMTLPTRELLALALA
AGLSLPAALAAEPFVASDIRVDGLQRIASGVTFYLPVNRGTVDGKVADSLRAL
YRTGFEDVQDRGNILIVTKRPAINKLTITGKNDIKSEELKGLGDIGLTFEGGT
FDRLSLDRVTELFRQYNNRKNVEITPTVPSLDRNRVDVAIAKEGKAARHVNAL
IGTEKFANEDILENKEHNWASVRRDQYSEKMSGDLKLSNWLDRGYDFSV
DSTQVSIQPKRDMFVSAGTEGQYKISEIKVGTDLVPOADVERLVPKAGDIFSR
ALLEVTSDAINTLSNGYAFKVNPIPTNREDTVAVNLQVVGPRVSRVILYKQ
NSRSTDEVLRERMPENTWYQAADRSKIRLQRLGCFESVDVETPAVPGSNQDQDV
VYNKETSISFVLGLFSQAFGVTTSVQLSQNNFLGGGNRVAVEARSTPQQRVAFS
YTNPFPTDGGVSLGYNLSWRLEDYSDFNATQFNSTNGSAQVVGCPITENDTVSAMIG

gene

complement(5091. .5918)
/gene="cdsA"
/note="XCC1368"
complement(5091. .5918)
/gene="cdsA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="phosphatidate cytidyltransferase"
/protein_id="AA040665.1"
/db_xref="GI:21112428"

CDS

complement(5091. .5918)
/gene="cdsA"
/note="XCC1368"
complement(5091. .5918)
/gene="cdsA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="undecaprenyl pyrophosphate synthetase"
/protein_id="AA040667.1"
/db_xref="GI:21112429"

VDSNQITTFPGTTPQAIVDYIQAIQTDTFTKAVRTFEGWARDTRNDFFMPTGRMYQRVG
LETTATLPGSTVEYKLYNQISNVPILIPALVNLTRFELGYGDSYKGRSGKTVTNADGTT
RPVATLPGTFYENFYAGTNSVRFENGTDLNPRSEAINGENRGPGLGSLKTVSGVEM
YFKLFDSPSARIASAFDFGNVYSDVDAFKANELRASTGVALLWRAPVGPISIVAFP
LKEDNDEIERLQTFGGQF"
complement(2525. .3871)
/gene="XCC1366"
complement(2525. .3871)
/gene="XCC1366"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="conserved hypothetical protein"
/protein_id="AA040664.1"
/db_xref="GI:21112426"
/translation="MGDFIGSVMMIVSLGVLVTFHERGFHWVARRCQVKVLRFSVGF
GKPLMRDRHGTEFATAAIPLGYSVXMLDEREGEVHPAEREQAFNRKTVQRTAIVA
AGPIANLLCLCHVLMWAMFVIGKDYISATVGRADGLAAAGLVPGERIVIRIDGRSVSSW
SDASMQLTAAAMDRDQVLTAAEEGNGNSEHTLRLSQLPAGDFERRVATLAGIGWQFM
LQPPVDKVVAGSAADGVLKPGDRIVAIDGQPIRSAGEVPAQLAQALQGTGGTGMIEVA
REDDRLALEIAPRAKSPGEGWMLGVGFATAAPAYDSROOYGAFAVPAARIRGTGKMTA
DSLGMKMLTGOASVKNSGPTVIARAANASAEKGVDMFLYELGLLSLALINLMP
IPLDGGHLLYLIELVKGSPISERAMIAGQYVGLAVLAGMLGAFYNDILGLVPR"
complement(3898. .5088)
/gene="dxr"
complement(3898. .5088)
/note="XCC1367"
complement(3898. .5088)
/gene="dxr"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AA040665.1"
/db_xref="GI:21112427"
/translation="MSGSSLRVAVFGATSGISALDVIARHPERLERASVLSAGSKV
DAALACVLRHPAHAVTADAALYPALRDGLRAAGLTTOAHAGDQALDAAASDACDTY
VAATVGAAGLSSTLAANAAGKRLILANKESLVLAGELITHTAAAGAGIITIDSEHSA
IFQCLRSCDASRGVRRVLTATAGSGPFRGRORAOIAEVTPAQAVHPKMGPKISVDS
ATLAKGLEIVIEAHLPLGCEOIDVLVHPQSLVSLVSEVDFVDSGLAQGLGIPDMRTTL
AVGLAWPVESEGVGGLDLQQLQGRDFEAPDTEAPFCLRLANDALRAGGTAPAILNAA
NEVAVSAFLQCKVGFLAIPALVEHTLTTLQRONADTLNALLFADAEARRTTERALAH
SLHA"
complement(5091. .5918)
/gene="cdsA"
/note="XCC1368"
complement(5091. .5918)
/gene="cdsA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="phosphatidate cytidyltransferase"
/protein_id="AA040666.1"
/db_xref="GI:21112428"
/translation="MTOTRVTAALIMAPLAICAIVLLPTOMLAALAILFTGLWEWL
KLGGIDSLPRTVLLMLLMVWASVGMVLFQIALVGVGVWLVALLWLRFFEN
FGADHGNQARALKLAAGTLAVLPAAALVLLHANPKGNLWLLTALTWVNAADSGAY
FAGRRFKHKLAPRISPNKTIAGGLVAGIACAFEGWAGLGLTAHVLPQVLVAAV
AVLASVIGDLFSELLKRRHAGKDSGTVPVGHGVDLRIDGVLAAALPVFALGKELIFG"
complement(5915. .6691)
/gene="upps"
complement(5915. .6691)
/note="XCC1369"
complement(5915. .6691)
/gene="upps"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/trans_table=11
/product="undecaprenyl pyrophosphate synthetase"
/protein_id="AA040667.1"
/db_xref="GI:21112429"

```
/translation="MASTAMHSEPSAAVPRHLAIIIMDGNRWAORRRPRVIGHRAG
ARAVNTIGDCLDKGVSALTLEAFSENNWQDEVDALMKLFLHALDRGVEELARRG
VQVRIIDGDSRETAFLRDMAGAEIRITLVAISAAISYGRDIDIAAARALAVEV
AAGRLQEOIDEALLASVALADPAPDLFRTGGDTRISNELLQWATLWTFETL
WPEFDAGVLQOALDDYAGRRERFGLTSAQIAEKATEASSA"
complement(6694..7251)
/gene="frf"
/note="XCC1370"
complement(6694..7251)
/gene="frf"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="ribosome recycling factor"
/protein_id="AAM40668.1"
/db_xref="GI:21112430"
/translation="MNLQIKDQAOITMTKSIDALRSLTITVETGRASPALLDNIKVA
YGTDLPLNOVASISVSEGRSLVIFDKGMIKDAVEKAIYASDLGTPTVGVIRLNL
PPLTERRKELSKSVHGEEDAKVAIRNIRRDANOQVKDILLKQVTEDEARGAEDDI
OKLTDKAIKDQDVVYKREQELMTV"
complement(7488..8210)
/gene="pyrH"
/note="XCC1371"
complement(7488..8210)
/gene="pyrH"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AAM40669.1"
/db_xref="GI:21112431"
/translation="MSELSYRILILKSGEALMGDGYCIDPKVINRLAHEVIERAQA
GAQVALVGGGIFRAGLAAGMDRVTDHGMGLATVINALAMODALEKLGAKVRVM
SAIKNDVCEDFIRRAIRHLEKRIAIFAAGTGNPFTDSDGALRALEIGADILLK
ATKVGVDKDPKKISDAVYDSLTLYDEVIMGLEVMTAFALARDSDLPLRFGMS
EPGLRLRHQAQIGTLVQGRS"
complement(8318..8983)
/gene="XCC1372"
complement(8318..8983)
/gene="XCC1372"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
Query Match 24.6%; Score 290.6; DB 1; Length 10029;
Best Local Similarity 57.2%; Pred. No. 3.2e-24;
Matches 612; Conservative 0; Mismatches 439; Indels 18; Gaps 4;
QY 2 TGGCAGCCTGTCGATCTTTGGGCGCACCGGCTCCATCGGCGAATCCACCTTCGACCTCG 61
Db 5072 TTGCTCGGGTCGCGGTGTTGCGGCGGACCGGCTCCATCGGCTCGGCACTGGAGGTGA 5013
QY 62 TCATCGCGAAGCGCGGCGCGGCTTCGCGACCTGCTGTCGACCGCGCGCGCAACA 121
Db 5012 TTGCGCGCA-----TCGGAGCGGCTGCGTCCAGCGTGTGTCGCGCGGCAAGG 4959
QY 122 TCGCGGCGACTGCGCGCAATGGCGTGCCTGAAGCGGAGTGTGCGTCAACGCGCATG 181
Db 4958 TCGATCGCTGCTGGCGTGTGCGTCTGCACCGCCCGCCGCGGCTGATCGCGGACG 4899
QY 182 AGAGCTGCTGCGCGGCTGCGGAGGCGCTG-----GCCGGAGCGGACCGAGGTG 235
Db 4898 CCGCGTGTATCGGCATTGCGCGATGTTGCGCGCGCGGCTGACCAACCGCGCGC 4839
QY 236 CGGCGCGGCGGCGGCGGCGCTGCTGCGGAGGCGCTG-----GCCGGAGCGGACCGAGGTG 292
Db 4838 ATCGCGTGTATCGGCGCTGGAGCGGCTGGCGGCGGAGCGCTGCGGACCGAGTGTG 4779
QY 293 CGGCCATCGTGGCGCGCGGCGGCTGCTGCGGAGTGGCGGCTGAAGACGCGGCGCGCA 352
Db 4778 CGCGGATCGTGGCGCGCGGCGGCTGCTGCCACCTGCGGCGCGGCGCGGCAACG 4719
```

```
QY 353 CGTGCGGCTGCGCCACAAAGAAAGCCTCGTGACGCGAGGCAACTCTCTGATCGGACGG 412
Db 4718 GCCTGTTGTCGCCAACAAAGAAATCCTGCTGTCGCGCGGAGCTGCTGATCCCGCACCG 4659
QY 413 CCAGGAGAACGCGCGCCACGATCCTCGCGTGGACAGGAGCACTCCGCGGTCTTTTCAGG 472
Db 4658 CGGCTGCGGCGCGCGGAGATCATTTCCGATCGACAGCGACAGTGGCGATCTTTTCAGT 4599
QY 473 CGCTG---GCGGCGGAGGACACGCGCTGCTGAGCGCGCTCATCATCACGCGCTCCGGCG 529
Db 4598 GCCTGCGCTCTCGACGCGCAGCGCGGCTGCGCGGCTGATCTCTACGCGCTCCGGCG 4539
QY 530 GCGCGTTCGCGGCTCGAGCCTCGAGCGATCCGCGCTGACCGGTGGCGGAGCGGAGG 589
Db 4538 GCGCGTTCGCGGCGCGCGCGTGCCTGCGCAATTGCGCGGAGTGCAGCGCGCGCAAGCGTGG 4479
QY 590 CCATCCCAACTGGTCCATGGCGCAGCGGATCTCCATCGACAGCGCTCGATGTTCAACA 649
Db 4478 CGCACCCGAATGGTGCATGGGCGGCGGAGATCTCGGTGCGACTCGGCGACGCTGATGAACA 4419
QY 650 AGCGCTCGAGCTGATCGAGACGCGCGAATTTCTTCGGCTTCGAGCGCGGACCGGATCGAGG 709
Db 4418 AGGCGCTGGAAGTATCGAGGCGCACCATCTATTGCGCTGCCCGGCGAGCAGATCGAGC 4359
QY 710 CGGTGCTCATCGCGAATCATCGTCATCGGATGTTGGCTTCTCGGAGCGGCGGCTGA 769
Db 4358 TGTGTTGTCACCGCAGAGCCTGTTGCTGCTGTTGGAAATTCGTGCGAGCGGCTCCACCT 4299
QY 770 TGGCCCATCTCGGCGCGCGGACATCGGCACGCCATCGGATTCGCGCTGAACTGCGCGG 829
Db 4298 TGGCCCATCTCGGCTACCGGACATCGGCACCGCTTGGCGTGGCGCTGGCGCTGGCGCG 4239
QY 830 GTGCGCGGAGTGTCGCGCTGCCCGGATCGACCTGCACAGATTTGGAGCGCTCACCTTCC 889
Db 4238 AGCGCTCGAGTCCGCGCTCGCGGCGCTGATCTGTCGAGCAAGCGCGGCTGGACTTCG 4179
QY 890 AGAGCTCGACGAGGAACGCTTTCGCGCGCTGAGGCTTGGCGGAGAGCTGATCGCGGCGG 949
Db 4178 AAGTCTCGGACACCGGAGGCGTTCCTCGTGTGCTGCGCTGGGAGCGCTTTCGCGCGCG 4119
QY 950 GGGCGCTGTCGGGCGCGCGCTTCAACGCGCGGAGAGATCGCGCTCGATCATTTTCATCG 1009
Db 4118 CGGCGACCGCGCGCGGCTCTGACGCGACCAACGAAGTGGCTGTTTCAGCGCTTCTTC 4059
QY 1010 CGGAGCATCGGGTTCGACATCGCGCGCTGTTGTCGAGGAGCGCT 1058
Db 4058 AGGCGAAGTGGTTCCTTAGCCATCCCTGCGTTCGTAACACAGCT 4010
RESULT 10
AF282879
LOCUS
DEFINITION
Pseudomonas aeruginosa 1191 bp DNA linear BCT 15-SEP-2000
reductoisomerase gene; complete cds.
ACCESSION
AF282879
VERSION
AF282879.1 GI:9664862
KEYWORDS
Pseudomonas aeruginosa.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
REFERENCE
1 (bases 1 to 1191)
AUTHORS
Altincicek,B., Hintz,M., Sanderbrand,S., Wiesner,J., Beck,E. and
Jomaa,H.
TITLE
Tools for discovery of inhibitors of the 1-deoxy-D-xylulose
5-phosphate (DXP) synthase and DXP reductoisomerase: an approach
with enzymes from the pathogenic bacterium Pseudomonas aeruginosa
FEMS Microbiol. Lett. 190 (2), 329-333 (2000)
JOURNAL
MEDLINE
20487113
PUBMED
11034300
REFERENCE
2 (bases 1 to 1191)
AUTHORS
Altincicek,B., Jomaa,H. and Beck,E.
TITLE
Direct Submission
```

JOURNAL Submitted (27-JUN-2000) JLU-Giessen, Biochemisches Institut,
Friedrichstr. 24, Giessen 35392, Germany
FEATURES
Source Location/Qualifiers
1. .1191
/organism="Pseudomonas aeruginosa"
/db_xref="taxon:287"
CDS
1. .1191
/function="biosynthesis of isopentenyl diphosphate, IPP"
/note="Dxp reductoisomerase; Dxr"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AAF97241.1"
/db_xref="GI:9664863"
/translation="MSRPQRIISVLGATSGISLTLDVVQVHDPDRYAEFALTGFSRLAE
LEALCLRHVPYAVVPEQAAIALGSLAAAGIRTVLFGEOALCEVASAPEDVMA
ALVAGALPSTLAAYEAGKRVLLANKALVMSGALFMQAVRSAGVALPIDSEHNAIF
QSLPRYADGLERVRRILLTASGPFRETLEQIALVYPEQACAHNWSMKRISV
DSAMNKGLEIEACWLFDAQPSQVYVHPSVHSMVDYVDSVIAQLGNPDHRT
PISYAMWPERIDSGVPLDMFAVGLDFORPEQRFCLRLASQAARETGSAPAMLN
AANEVAAFLERHFRISDIAVIEDVLNREAVTAVESLDQVLAADRARSVAGQWL
RHAG"

BASE COUNT 176 a 400 c 413 g 202 t

ORIGIN

Query Match 24.2%; Score 286; DB 1; Length 1191;
Best Local Similarity 55.8%; Pred. No. 1.9e-23;
Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY 20 TTGGGGCCACCGCTCCATCGGCGAATCCACCTTCGACCTGCTGCGGAAGGGCG--- 76
DB 29 TCGGCGCGACCGCTGATCGGCTGAGCACCTGACGCTGCTCCACGCTCATCCGATC 88
QY 77 GGCCCGAGGGTTCCGACCGTCTGACGGCGGGCGACATCCCGCGACTGCGCG 136
DB 89 GTTACGAAGCTTTCGCGCTGACTGGCTTCAGCGCGCTGCGCGAATCGAGCGCTGTGCC 148
QY 137 AATGCGCGTGGCTGAAGCGGCGAGCTTCCGCTCACCGCGATGAGGACTGCTCCCGC 196
DB 149 TCAGGCACCGCCCGCTTATGCGGTGTCGCGGAGCAGCGCGCGGCGATTGCTTCGAG 208
QY 197 CGTGGCGAGGGCTGGCGGCGACCGGACCGAGTTCGCGGCGCGCGAGCCATCG 256
DB 209 GCTGCTGCGCGCGGCGGTATCCGCCCGCGGTGCTGTTCGCGAGCAGCGCTTCGCG 268
QY 257 CCGAGCGCGCGACCGCGCGGCGACTGACCATGTCGCGCATCTGCGCGCGCGCGCGC 316
DB 269 AAGTGGCAGCGCGCGGAGTGACATGTTATGCGGCGCATCTGCGCGCGCGCGCGC 328
QY 317 TCGTGGCGGAGTGGCGGCGCTGAAGCAGCGCGCGCGCATGCTGCGCGCTGCGCAACAGGAA 376
DB 329 TGCGGTGACCTTGGCGGCGCTGAGCGCGCGCGCGCGCTGCTGCGCGCAACAGGAGG 388
QY 377 GCTGCTGAGCGGACGACTCTCTATGCGGAGCGCGCGCGCGCGCGCGCGCGCGCATCC 436
DB 389 CGTGTGTGATGTCGCGGCGCTGTTTCATGAGCGCGCTGATGAGCGCGCGCGCGGCTCTC 448
QY 437 TGCGGTGAGCAGCAGCAGCTCGCGGCTCTTTCAGCGCGTGGCG-----G 481
DB 449 TGCGGATGACACGAGCAGCAGCATCTTCAGTCTGCTGCGCGCGCATTTATCGCGATG 508
QY 482 GCAGGACAGCGCTCGGTGAGCGGCTCATCATCAGCGCGTCCGCGGCGCGCTTCGCGG 541
DB 509 GCCTGAGCGGGTGGCGGTGGCGGATCTCTTTCAGCGCGCTCCGCGCGCGCTTCGCGG 568
QY 542 ACTGGAGCTCGAGCGCATCCGCGCTGACCGTGGCGCGCGCGCGCGCGCGCGCGCAACT 601
DB 569 AGACGCGCTGGAGCAACTCGCTTCGCGTACCGCGCGCGCGCGCGCGCGCGCGCAACT 628
QY 602 GGTCCATGGCGGAGGATCTCCATGACAGCGCGCTTCGATGTTCAACAAGCGCGCTCGAG 661
DB 629 GGTGATGGGCGTAGATTTTCGTCGACTCCCGCAGCATGATGAACAAGGGGCTCGAAC 698

QY 662 TGATCGAGAGCGCGCAATTTCTCGGTTTCGAGCGGACCGGATCGAGCGGTGCTGCATC 721
DB 689 TGATCGAGCGGTGCTGGCTTTCGAGCGGACCGGAGGTGAGGTGATGATCCACC 748
QY 722 CGCAATCCATCGTCCATCGATGGTGGCTTTCGAGCGGCGGCGCTGATGCGCCATCTCG 781
DB 749 CGCAGAGCGTGTATCCACTCGATGGTGGACTACGTCAGCGGTTCGGTATGCGCAGCTCG 808
QY 782 GCCCGCGCGACATCGCGCCACCGCATCGGATTCGCGTGAACCTGCGGGTTCGGGGGAGG 841
DB 809 GCAATCGGACATCGCGACCGCATTTCTATGCCATGGCTGCGCGGAGGATCGATT 868
QY 842 TGCCCGTCCCGCGATCGACCTCGACAGATTCGAGCGCTCACCTTCCAGAACGCTGACG 901
DB 869 CCGCGCTTTCGCGCGTGGATATTTTCGCGTTCGCTGCTGATTTCCAGCGCGCGGACG 928
QY 902 AGGAACGCTTTCGCGCGCTGAGCTTTCGCGGAGACGCTCATGCGCGCGCGCGCTGTCG 961
DB 929 AGCAGCGCTTCCCTCGCTGCGCTGCGAGCGCGCGGAAACCGCGCGCGCGCGCGCC 988
QY 962 GCGCGCGCTTCAACGCGCGCAAGGAGATCGCTCGATCATTTTCATCGCGCGGACGATCG 1021
DB 989 CGGCGATGCTGAATGCGCGGACGAGTGGCGCTGCGCGCATTTCTCGAGCGGCATCC 1048
QY 1022 GGTTCCTGACATGCGCGCGGTGCTCGAGGAGCGCTCGCGCGGCTTCG 1071
DB 1049 GCTTCAGCGACATCGCGGTTCATCATCGAGGACGCTGTAACCGCGGAGCG 1098

RESULT 11
LOCUS AE004785/1 14537 bp DNA linear BCT 30-AUG-2000
DEFINITION Pseudomonas aeruginosa PA01, section 346 of 529 of the complete genome.
ACCESSION AE004785
VERSION AE004785.1 GI:9949809
SOURCE Pseudomonas aeruginosa.
ORGANISM Pseudomonas aeruginosa.
REFERENCE 1 (bases 1 to 14537)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
TITLE Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
JOURNAL Nature 406 (6799), 959-964 (2000)
MEDLINE 20437337
PUBMED 10984043
REFERENCE 2 (bases 1 to 14537)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z., Paulsen,I.T., Reizer,J., Saier,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA

FEATURES
source Location/Qualifiers
1. .14537
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
gene complement(66. .1418)
/gene="PA3649"
CDS complement(66. .1418)
/gene="PA3649"

```
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG07037.1"
/db_xref="GI:9949810"
/translation="MSALNMIYVGTVALGVLTFFHEFFHWARRCGVVKVLRFSVCGF
GPLVRHHRHGFVFAAIPLGYYVKMLDREAEPYPAHLLEQSFNFKTVRQRIAIVAA
TPIANFLLAIFFWVVALGSOQVRVIGSAPEASLAQAQLEAGQELLAVDGEPTVG
WNGVNLQVLRRLGESQTLFVRVQEKGSNDVSTHOVRLDGWLKGEDNPDPIASLIGIRPW
RPALPVLAELOPKGPAAGLKLGRLOS IDGIAVDDMQVVDSVRARPGORVOLKV
LRDGEVDVALAVRGECKARSGYMGAGVAGTEWPAEMLREVSYGPLEAVGOALSRT
WTMSLTLDSIKLMSLSESVKNSLSPITIAKVAGASQSGVGFNFLAYLSISLGV
LNLPLIPVLDGGHLLPYLVEWVRGRPLSERVQAWGMQIGISLVVGVVMLLALVNDLSRL
"
complement(1444..2634)
/gene="dxx"
/notes="PA3650"
complement(1444..2634)
/gene="dxx"
/codon_start=1
/transl_table=11
/product="1-deoxy-d-xylulose 5-phosphate reductoisomerase"
/protein_id="AAG07038.1"
/db_xref="GI:9949811"
/translation="MSRPQISVLGATSGISLTLDVQRHDPDRYEPALTFGSRLAE
LEALCULRHPVYAVVPEQAAIALOGSLAAAGIRTVLFGEQALCEVASAPEVDMYMA
ATVGAAGLSTLAAGVAGKRVLLANKEALVMGALFMQAVKRGSGAVLLPIDSEHNAIF
QSLPRNADGLEVRGVRILLTAAGGPFRETPLEQLASVTPEQACAHPNQSMGRKISV
DSAMNKGLELLEACWLPDAOPSOVEVTHPQSVIHSMDVYVDGVIQALGNPDMRT
PISYANWPERIDSGVPLDMFAVGRDLDFQDEQREPCRLASQAQETGGSPAPMLN
AANEVAAFLERHIAFRFSDIAVTIEDVLNREAVTAVESLDQVLAADRARSVAGOWLT
RHAG"
complement(2631..3446)
/gene="cdsA"
/notes="PA3651"
complement(2631..3446)
/gene="cdsA"
/codon_start=1
/transl_table=11
/product="phosphatidate cytidyltransferase"
/protein_id="AAG07039.1"
/db_xref="GI:9949812"
/translation="MLKQRIITALLPITAGLGGFFLEGAFFALFICAVVSLGAWETA
RLAGYEQQGRVAYATVAVLMVALYHLPOLAGAVILLALVMTLATVILVITYPSVG
YMGRRVRLGMLGILLIPAWQGLVLLKQWPLANGLIAYVVLVWGADIGAYFGSKAFG
KRKLAPRVSPKGSWEVYGGLAASLAIITLAVGLYRGWSGLALLLALIGALVVFVSIV
GDITSMFKRQSGIKDSSNLLPCHGGVLDRLDSLTAAIPVFAALLWAAGWGA"
complement(3440..4195)
/gene="upps"
/notes="PA3652"
complement(3440..4195)
/gene="upps"
/codon_start=1
/transl_table=11
/product="undecaprenyl pyrophosphate synthetase"
/protein_id="AAG07040.1"
/db_xref="GI:9949813"
/translation="MEKTRKDVCPVRHVAIIMDGNNRWAKRLLPLPGVAGHKAGVDAR
AVTEVCAAGVEVLTLLFAFSENQRPADVSALMELFLVALRREVRLKLDENGIRLRI
IGDRTFPHPELOAMBEAATAAGTFRLLQVAANYGGOWDIIQAAQRLAREVOGGHL
AADDISALLQGLCVTGDQPLDLCINTGGEHRI SNFLLWQLAYAEIYFSDLEFWPDFK
HAAMRAALADFSSKRQRRFGKTSQSEVAEARPSC"
complement(4211..4768)
/gene="frr"
/notes="PA3653"
complement(4211..4768)
/gene="frr"
/codon_start=1
/transl_table=11
/product="ribosome recycling factor"
/protein_id="AAG07041.1"
/db_xref="GI:9949814"
/translation="MINEIKEAQERMGKLTLEALGHAFKIRTGRAHPSLDSVMVSY
YCADTPLROVANVTVEDSRTLALAVFDKSMIQAVKAIMTSDLGSLNPATAGTTIRVPM
PALTEETKRGYTKQARAEEQARVSVNRIRRALAQKLDQKEISEDEERRAGDOV
QKLTDRFGEIEKALEAKEADLMAY"
complement(4771..5508)
/gene="pyrH"
/notes="PA3654"
complement(4771..5508)
/gene="pyrH"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AAG07042.1"
/db_xref="GI:9949815"
/translation="MAQQLSARQPRYKRILLKLSGALMGSEEFIDPKVLDMALEI
GQLVGIGVQGLVIGGNLFRGAALSAAGMDRVTDGDMGLMGLATVMNGLAMRDALERSN
IPALVMSAISWVGVTQDHYDRRKAMRHVGGEVYVFSAGTGNPFTTDSACLRAI EID
ADVLLKATKVDGYVTADPFKDPNAEKFERLTDEYVLDRLKLVMDLTAICLCRDQNMPL
RVFNMKPGALLNIIVGGAEGTLIEG"
complement(5707..6576)
/gene="tsf"
/notes="PA3655"
complement(5707..6576)
/gene="tsf"
/codon_start=1
/transl_table=11
/product="elongation factor Ts"
/protein_id="AAG07043.1"
/db_xref="GI:9949816"
/translation="MAEITAAVMVKELRERTGLGMMECKKALTAAAGDIEKAIDDMRAA
GAIKAAKAGNTAAEGSIAVKIAADNKAAYIIEVNSOTDFLALQDDQFKGFAESLEKA
FNEKLTDAAPLYEAREEARLALVAKTGENVIRLRLTRVSGDVVGAYLVHGRIGVYVNL
KGNPBLAKDIAHVAASNPQFLSASEVSEETAKEIFEFLALNADKIAKGKAPENIVEN
MYKRISKLAEALSVLEQPFVKNPQVKGDLAKQAGAEIVSFVRYEVEGEGIAEVDFF
AAEVAQAQVAATQ"
complement(6707..7447)
/gene="rpsB"
/notes="PA3656"
complement(6707..7447)
/gene="rpsB"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S2"
/protein_id="AAG07044.1"
/db_xref="GI:9949817"
/translation="MSQVNRMDLKGAVHFGHGTQRYWNPKNMGKFIIGARNKIHIINLE
KITLPMEALTYFVERLAAGKNKILFVGTAKSAGKIVREAAACGMPYVDHRWGLGMLT
NFKTIRQSKIRLRDLETQSDQGFDKLTKEALMRSDLEKLEKRSGLGKIDMGGLPDA
LYTVIDVHERIAITEANKLGIPIGVVDTNNSPESGYVYVPIGNDDAIRAQVLYLNSWA
EAVIRKQGNATSADEFVEEAPAESAG"
7712..8497
/gene="map"
/notes="PA3657"
7712..8497
/gene="map"
/codon_start=1
/transl_table=11
/product="methionine aminopeptidase"
/protein_id="AAG07045.1"
/db_xref="GI:9949818"
/translation="MTVTIKTPDDIEKMRAGRLAAAEVLEMIGEHKPGVYVTEELDRI
CHDYIVNEOKAIPAPLNYKGFPSICTSINHVVCHGIPNEKPLKEGDLNVDITTVIKD
GYHGDTSKFLVKGTPEWADRLCQITQECNYKGISVVRPCAHLGIDGIEIIOKHAEKNG
FSVRYECGHGIGKVFHEEPQVILYGRAGTGIELKGOMITFIEPMINOGRPETRLIGD
GWTAITKDRKLSAQWEHTVLVTDAGYEIILFLRNDFTFPRTSAA"
8664..11366
/gene="glnD"
/notes="PA3658"
8664..11366
/gene="glnD"
/codon_start=1
/transl_table=11
/product="protein-PII uridylyltransferase"
/protein_id="AAG07046.1"
```



```

/db_xref="gi:9949819"
/translation="MPQVDFPDRGQFAELALKSP IAAFKKAIROFREVLDNRFN
SGRIRLIEDRAWVDIQLQOAWORFDDADIALVAVGVYGRGLHPYSDVDLLI
LLDEDSQEPREPTEGFTLLWDIGLEVGSVRVQOQAEARADLVITWMECRTI
CGPDLRQMLQVTSAAHWPSEKFFLAKRHEQORRAKYNDEYNPNVKGSPGGL
RDQITILMARQVGSNLHVALVREGFLVESEKMLASSQEFILMRVYLAHMLAGRAE
DRLLDFHORSIARLFGEDNDVKLAVERFMOKYRVVMAISELNDLIQHFEFVILPC
FOVQIOLPNSRFLGRDGYIEVHPNVFKTFPALLIEFVLMQHPKGYRADTIRL
LRDSRHLLIDDFRHDIRNTSLFIELFSSQCIHRNLRMRYGILRPLPFGHIGQ
MOHDLFTYTVFDTHTLNLIKHLRKNRMAEKYPLASKIIDRLPKPELIYAGLYHD
IAKGRGDHSELGAVDAEAFCSHQLPLWDRQLVSLVQNLHVMSTTAQRKDLSDPOV
IFDPAQLVGDGTHLDYLVLTADINATNPTLWNSWRASLLRQLYTETKRALRGLN
Query Match      24.2%; Score 286; DB 1; Length 14537;
Best Local Similarity 55.8%; Pred. No. 9.4e-24;
Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY 20 TTGGGGCCACCGCTCCATCGCGGAATCCACCTTCGACCTGTCATGCGGAAGGCGG--- 76
DB 2606 TCGGGCGGACCGGCTCGATCGGCCTGAGCACCTCGAGCGTCTCGCAGCGTCATCCCGATC 2547
QY 77 GCGCGGAGCGCTTCGCGACCGCTGCTCACCGGCGGCGCAACATCGGCGGACCTGGCGG 136
DB 2546 GTTAGAGCCTTCGCCCTGACTGCTTCAGCGCGCTGCGCGAATCGAGGCGCTGTGCC 2487
QY 137 AAATGGCGCTGCGTGAAGCGGAGCTTGGCGTCCACCGCATGAGGACTGCTTGGCGCG 196
DB 2486 TCAGGACACCGCGCTATGCGGGTGGTCCGAGCAGCGCGCGGATGCTTGGCAGG 2427
QY 197 CGCTGGCGAGCGCTGCGCGGAGCGGACGAGGTCGCGGGCGGGCGGAGGCCATCG 256
DB 2426 GCTCGCTGCGCGCGCGGGTATCCGACCCGCGGCTGTCTGTCGCGGAGCAGCGGTTGTGCG 2367
QY 257 CCGAGGCGCGCGGCGCGGCGGCGGACTGGACCATGTCGGCCATGTCGGCGCGCGCGGCG 316
DB 2366 AAGTGCCAGCGCGCGGCGGAGTGGACATGTAATGGCGGCATGTCGGCGCGCGCGGCG 2307
QY 317 TCGTGGCGGGAATGCGGCGCTGAAGCAGCGGCGGCGGCGGCTGCGGTCGCGCAACAGGAA 376
DB 2306 TGCCGTCGACCTGCGCGCGCTGAGGCGCGGCGGCAAGCGGCTACTGTCGCGCAACAGGAG 2247
QY 377 GCTGTGACGCGGCGGCAACTCTGTATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGATCC 436
DB 2246 CGCTGTGATGTCGCGCGCGCTGTCATGCGAGCGGCTCAAGCGGCGGCGGCTGCTCC 2187
QY 437 TGCGGTCGACAGCAGCAGCTCCGCGGCTTTTCAGCGCGCTGCGG-----G 481
DB 2186 TGCGGATCGACAGCAGCAGCAGCTTCCAGTCTGCTGCGCGCAATATGCGGATG 2127
QY 482 GCGAGACACGCGCTGCGTGTGAGCGGCTATCATCATCGCGGCTGCGGCGGCGGCTTCCGCG 541
DB 2126 GCGTGGAGCGGTCGCGGCTGCGCGGATCTCTTGACCGCTTCCGCGCGGCGGCTTCCGCG 2067
QY 542 ACTGGAGCTTCAGCGCATCCGCGCTGACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCTCC 601
DB 2066 AGACGCGGTCGAGCAACTCGCTTGGTGAACCGCGGAGCGGCTGTGCGCACCGCAACT 2007
QY 602 GGTCTATGGCGGCGGATCTCCATCGACAGCGCTCTGATGTTCAACAGGCGCTCGAGC 661
DB 2006 GGTGATGGGCGTAAATTTCCGTCGACTCGCGCAGCATGATGAACAGGCGCTCGAC 1947
QY 662 TGATCGAGCGCGGAATTTCTTGGCTTGCAGCGCGGAGCGGATCGAGGCGGCTGCTCCATC 721
DB 1946 TGATCGAGCGGTCGCTGCTGCTGCGAGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCT 1887
QY 722 CCAATTCATCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
DB 1886 CCGAGGCGTATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1827
QY 782 GCGCGCGCGGACATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
DB 1826 GCAATCCGAGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1767

```

```

QY 842 TGCCCGTCCGCGGATCGACCTCGCACAGATTTCGAGCGCTCACCTTCACAGAGCCTGACG 901
DB 1766 CCGCGGTTTCGCGCGCTGGATATGTTCCGCGTCTGCGCTGGATTTCACAGCCCGGACG 1707
QY 902 AGGAACGCTTTCCGCGCTGAGCTTTCGCGGAGAGCTCATGCGGCGGCGGCGGCTGCTCGG 961
DB 1706 ACAGCGCTTCCTCCCTGCTGCGCTTGGCGAGCAGCGCGCGGAAACCGCGCGGCGCGCC 1647
QY 962 GCGCGCTTCAACGCGGCAAGGAGATCGCGCTCGATCATTTTCATCTCGCGGAGCGATCG 1021
DB 1646 CGCCATGCTGAATGCGCGGACAGAGTGGCGTGGCGGCGATTTCTCGAGCGGCATCC 1587
QY 1022 GCTTTCGACATGCGCGGCTGTCGAGGAGCAGCGCTCGCGGCGGCTTCG 1071
DB 1586 GCTTCAGCGACATCGCGGTTATCATCGAGGAGCTGTGAACCGCGGAGCG 1537

RESULT 12
SC5H4 Streptomyces coelicolor cosmid 5H4. linear BCT 12-MAY-2002
DEFINITION Streptomyces coelicolor A3(2).
ACCESSION AL355913.2 AL645882
VERSION AL355913.2 GI:20520782
KEYWORDS 1-deoxy-D-xylulose 5-phosphate reductoisomerase; acyl CoA
dehydrogenase; aldehyde dehydrogenase; ATP/GTP binding protein;
beta-galactosidase; cytidine deaminase; gcpE; glycosyl transferase;
integral membrane protein; kinase; lipoprotein; metalloproteinase;
prolyl tRNA synthetase; regulatory protein; secreted sugar
hydrolase; sugar hydrolase; transcriptional regulator;
transcriptional termination/antitermination factor; translation
initiation factor IF-2; transport system integral membrane protein;
two-component system response regulator; two-component system
sensor kinase.
SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 39744)
AUTHORS Redenbach,M., Kieser,H.M., Denapait,D., Eichner,A., Cullum,J.,
Kinashi,H. and Hopwood,D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE 97000351
PUBMED 8843436
REFERENCE 2 (bases 1 to 39744)
AUTHORS Murphy,L. and Harris,D.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 39744)
AUTHORS Cerdano,A.M., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT
Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
numbered using the following system eg SC7B7.01c. SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous. The length
in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon. Gene

```

prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krohn et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 5H4 Lies between and overlaps cosmids 8B7 and 9F2 on the AseI-B genomic restriction fragment.

FEATURES

source Location/Qualifiers
 1. .39744
 /organism="Streptomyces coelicolor A3(2)"
 /strain="A3(2)"
 /db_xref="taxon:100226"
 /clone="cosmid 5H4"
 complement(2..136)
 /notes="nominal overlap with Streptomyces coelicolor cosmid SC8B7"
 misc_feature
 689..836
 /notes="nominal overlap with cosmid St5H4 from 1 to 148"
 gene
 1219..2877
 /gene="SC5H4.02"
 /note="SCO5678"
 1219..2877
 /gene="SC5H4.02"
 /note="SC5H4.02, possible regulatory protein, len: 552 aa; similar to TR:Q9ZBP0 (EMBL:AL034492) Streptomycetes coelicolor putative regulatory protein SC6C5.05c, 569 aa; fasta scores: opt: 310 z-score: 323.3 E(): 1.5e-10; 30.0% identity in 587 aa overlap"
 /codon_start=1
 /transl_table=11
 /product="putative regulatory protein"
 /protein_id="CAB91114.1"
 /db_xref="GI:7801250"
 /db_xref="SPTREMBL:Q9KYT7"
 /translation="MPPTLASLVHHSALKLTVRAGEDRLDVPVWVAHSELADPPVPM EGELLITALLKIDEDPEAMRYVVRKLAGVVGGLGFAVGVNDVPEALVDAAQOE GLPLEVPRPTPELTSKAVSAIAADQAVTAGFAOQRELTRRTLTGPGELLAAL AQVDCWALYDASGAVASAPEWAGRRARLTADYORLRERAPASSVVGAGNTGH SENADRVLSLGTSSRRPRAALAVGTAAALGTAERYAVHSAIALTLTTERNSLHEA GLKIDAAVRLMLLAGEPDHARTVAGDLYGLLDADPRVIVAESGAARNRGPTAAHTAA GTGGDPGGDLPTGLSLAAEASAAAGEANLAVPDGERLVVLPATDGGAAVACVHEHSALEAARPAEPLAGGDEESLVVGSAPSGPTAAASAAKQAQALSVAARRGRVCVE HEHVAAGSVLPILLADDAVRAFADGLRLRDHDATGCGDLVASVRAWLSRHGQWDAADLGVHRHTLVYRMRRVEELGRSLDDPDVRLWLAKATAPE"
 3021..4466
 /gene="SC5H4.03"
 /note="SCO5679"
 3021..4466
 /gene="SC5H4.03"
 /note="SC5H4.03, probable aldehyde dehydrogenase, len: 481 aa; similar to TR:Q66573 (EMBL:AE00677) Aquifex aeolicus aldehyde dehydrogenase AldH1, 476 aa; fasta scores: opt: 1129 z-score: 1260.2 E(): 0; 39.9% identity in 471 aa overlap and to SW:GAPN_STRMU (EMBL:L38521) Streptococcus mutans NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9) GapN, 475 aa; fasta scores: opt: 893 z-score: 997.5 E(): 0; 35.4% identity in 477 aa overlap. Contains Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenase family"
 /codon_start=1
 /transl_table=11
 /product="putative aldehyde dehydrogenase"
 /protein_id="CAB91115.1"

/db_xref="GI:7801251"
 /db_xref="SPTREMBL:Q9KYT6"
 /translation="MTSTHAFWLAGQAAGESGFDVTSPMDGTTTGVTSPLTDAQVEE AVAAAVAYRDEFAATPAHVRAAALDHVGRLEARTETRIARLISAENGKPKVMARGEVG RAVSVFRFAEARENRNGEAQRLTDAGQGQGLAETLDRPPKGVVIGIAPFNPPLNC AHKIAIPAAGALILKIPATPLSLGGLIAGELAEITLDPAGSVLILPVPNKPMPALVQ DELPLVISTGSETSVGYIADMSVPRKHCTLELGGNGAAVVLADMSDEDIDRAATRIA TFSNYGGQCSISVGYIADMSVPRKHCTLELGGNGAAVVLADMSDEDIDRAATRIA AERVEAWREAVDAGAKLTGGKDGASVAPTVLTLPADTTLAHEEVFGPVLVSQVRV TGPAEFAAVANSKSYGLOAGVTFHDLQAAFRAHRALEVGVGWIGIDIPSYRADQMPYGG AKQSGVGRGVKPFAMDDYYERVVLVTGLAL"
 3048..4454
 /gene="SC5H4.03"
 /note="Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenase family, score 357.30, E-value 1.6e-103"
 complement(4476..4889)
 /gene="SC5H4.04c"
 /note="SCO5680"
 complement(4476..4889)
 /gene="SC5H4.04c"
 /note="SC5H4.04c, probable cytidine deaminase, len: 137 aa; similar to TR:Q53367 (EMBL:AL021841) Mycobacterium tuberculosis cytidine deaminase Cdd, 133 aa; fasta scores: opt 185 z-score: 244.6 E(): 3.6e-06; 37.4% identity in 115 aa overlap and to SW:CDD_MOUSE (EMBL:AA388666) Mus musculus cytidine deaminase (pc 3.5.4.5) Cdd, 146 aa; fasta scores: opt: 173 z-score: 229.1 E(): 2.6e-05; 29.2% identity in 113 aa overlap. Contains Pfam match to entry PF00383 dCMP_cyt_deam, Cytidine and deoxycytidylate deaminase zinc-binding region and match to Prosite entry PS00903 Cytidine and deoxycytidylate deaminases zinc-binding region signature"
 /codon_start=1
 /transl_table=11
 /product="putative cytidine deaminase"
 /protein_id="CAB91116.1"
 /db_xref="GI:7801252"
 /db_xref="SPTREMBL:Q9KYT5"
 /translation="MTSOTNPVDHELIEAAAHVARTRCRGNHTMAAAARARDGRIVT AVNAVHTGTPCAELVVGAAAGVVELETIVAVGDRGVVPPCGCRQCLVDYFP GLEVIIVGEGDRVAPVPTDLLPETVYWAHDQDLAG"
 complement(4575..4820)
 /gene="SC5H4.04c"
 /note="Pfam match to entry PF00383 dCMP_cyt_deam, Cytidine and deoxycytidylate deaminase zinc-binding region, score 10.30, E-value 0.027"
 complement(4599..4724)
 /gene="SC5H4.04c"
 /note="PS00903 Cytidine and deoxycytidylate deaminases zinc-binding region signature"
 complement(4915..6090)
 /gene="SC5H4.05c"
 /note="SCO5681"
 complement(4915..6090)
 /gene="SC5H4.05c"
 /note="SC5H4.05c, possible glycosyl transferase, len: 391 aa; similar to TR:Q9X7Z1 (EMBL:AL049497) Streptomyces coelicolor putative glycosyl transferase SC6G10.05c, 412 aa; fasta scores: opt: 1364 z-score: 1557.1 E(): 0; 56.4% identity in 385 aa overlap. Contains Pfam match to entry PF00534 Glycos_transf_1, Glycosyl transferases group 1"
 /codon_start=1
 /transl_table=11
 /product="putative glycosyl transferase"
 /protein_id="CAB91117.1"
 /db_xref="GI:7801253"
 /db_xref="SPTREMBL:Q9KYT4"
 /translation="MSDSRARTLIVTNDPFRQGGIETFTVRELADRPDPDEVVLFS PTAAPGSGESFPYVVRHPVTRLLPTTGATAHAAAVARRYGDVWFGAAAPLEAMAG RLRTAGIGTAVASTHGVHWNWARTPGRRLLRRIGSQVDTLTLGERTRAPTEALAA AGTRTARLPVGDVTGTFHPAADGTRVTRYGLGHRPVLCAARLVPRKQDQTLRALP WYRRAPDAVILLVCGDPAAGLROLAREGVWDSVHFAGGHPHALLPAFYAADAFAP MPCRTRKGLGLEGLGIVYLEAASGLPVVAGDSGAPDAVREGTGHVVDGKSVANT

Query Match 23.6%; Score 278.6; DB 1; Length 39744; Best Local Similarity 56.1%; Pred. No. 4.7e-23; Matches 602; Conservative 0; Mismatches 444; Indels 27; Gaps 3;									
QY	9	CTGTGTCATCTTTGGGCGCACCGGCTTCATCGGGAATCCACTTCGACCTCGCTCATGGG	68						
DB	24771	CGTGGTGGTCTTCGGGTCCACCGGTGATCGGACCCAGGCGCATGACCTCGTGTGTCGG	24830						
QY	69	GAAGGCGGGCGCGAGCGCTTCGCGACCGTCTCTGACCGGGCGGGCGCAACATCCGGCG	128						
DB	24831	CACCCGGACCGGTTCCGGGTACCGCCCTGTCCGCCAAGCGGCGCGGGTCCGCCCTC--	24888						
QY	129	ACTGGCCGAAATGCGCGTGGCTGAAGCGGAGCTTTCCTGTACCGGCGCATGAGGACTG	188						
DB	24889	-CTCGCGGACGCGGTACCGGCTGAAGGCGCGGACCGTTCGCTCGCCCGCGGAGGACGT	24947						
QY	189	CTGTCGCCGCTGCGGAGCGCT-----GGCCGGGACGGGACCGCA	230						
DB	24948	CGTGGCGGCGCTCCGGAGGCGCTCACGCCACGTACGGCACGGGCGAGCGCTCCCGGA	25007						
QY	231	GFTCGCGGGGCGGCGAGGCCATCGCGAGGCGCGCGACCGGGCGGCGGCGGCGGCGG	290						
DB	25008	GATCCTCGCGGACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	25067						
QY	291	GTGCGGCATCTGTGGGCGCGGCGGCTCTGTCGCCGGAATGCGGCGCTGAAGCACGGCG	350						
DB	25068	CACGGGCATCACCGGCTCCATCGGACTCGGCGGCGGCGGCGGCGGCGGCGGCGG	25127						
QY	351	CAGCTGGGCTCGCCCAACGAAGAAAGCCCTGTCGACGCGGAGGCAACTCTGTATCGG	410						
DB	25128	CACCTCGGCTGCGCAACGAAGAGTCGCTCATGTCGGGCGGCGGCTGGTCAAGGCGCT	25187						
QY	411	GGCCAGGAGAAAGCGGCGGCGGCGGCTTCGCGGTGGACAGGAGCACTCCGCGGCTTTC	470						
DB	25188	GGCCAGGCG-----GGCCAGATCATCCGCGTGCATCGAGACGCGGCGGCTTCCA	25241						
QY	471	GGCGCTGGCGGGGAGGACACGGCTCGCTCGAGCGGCTCATCATACGGGCGTCGGCGG	530						
DB	25242	GGCGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	25301						
QY	531	GCGCTTCGGGCGCTGAGCGCTCGAGCGCATCCGCGCTCGACCGTGGCGGCGGCGGCG	590						
DB	25302	CCGCTTCCGGGGCGGTACAGGACGAGCTGGCGGCGCTCACCGTCGAGGACGCGCTCG	25361						
QY	591	CCATCCCACTGTCTCATGGCGGCGGATCTCATCGACAGCGGCTCGATGTTCAACAA	650						
DB	25362	CCACCCCACTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	25421						
QY	651	GGCGCTCGAGCTGATCGAGCGGCGGAAATCTTCGCGCTTCGAGCGGACCGGATCGAG	710						
DB	25422	GGCGCTGGAGGTGATCGAGCGGCGGCTCTAGGACATCCCTTCGACCGCATTGAGGT	25481						
QY	711	GGTGTGTCATCCGAATCCATCGTCCATGGATGGTGGGCTTCTCGGACGGGCGGCTGAT	770						
DB	25482	GGTGTGTCACCGCGAGTGTATGTCCACTCGATGTTGAGTACAGCGGAGCATCGACCT	25541						
QY	771	GGCCATCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	830						
DB	25542	GGCCCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	25601						
QY	831	TCGCGCGAGGTGCGGCTGCGGCGGATCGACCTCGCACAGATTCGAGCGGCTCACTTCCA	890						
DB	25602	ACGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	25661						
QY	891	GAAGCTCGAGGAGACGCTTTCCGGCGGCTTGAGGCTTTCGCGGAGACGCTCATGGCGG	950						
DB	25662	CCGCTCGACAAAGGCGCTTCCCTCCGTCGACCTGGCCAGGCGACGTCGCGGAGCTCG	25721						
QY	951	CGGCTGTGCGGCGGCGGCTTCAAGCGGCGGCGGAGAGATCGCGCTCGATCATTTTCGCG	1010						
DB	25722	GGGCGGCGGCGGCGGCTTCAATGCGGCGGCGGAGGAGTTCGCTCGAGGCGGCTTCGCT	25781						

QY	1011	CGGACGATCGGCTTCTGACATGGCGGCGGCTGCTGCGGAGGACGCTCGCGG	1063
DB	25782	CGGCGGCTGCGGCTTCTGCGGATCATGGAGACCGTCAACGCGGCTGCTCGAGG	25834
RESULT	13		
AE011773/c			
LOCUS			
DEFINITION	Xanthomonas axonopodis pv. citri str. 306. linear BCT 29-MAY-2002 the complete genome.		
ACCESSION	AE011773		
VERSION	AE011773.1		
KEYWORDS	GI:21107580		
SOURCE	Xanthomonas axonopodis pv. citri str. 306.		
ORGANISM	Xanthomonas axonopodis pv. citri str. 306. Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xanthomonas.		
REFERENCE	1 (bases 1 to 10242)		
AUTHORS	da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorfy, H., Faria, J.B., Ferreira, A.C.S., Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira, J.R.H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezra, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.		
TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities		
JOURNAL	Nature 417 (6887), 459-463 (2002)		
MEDLINE	22022145		
PUBMED	12024217		
REFERENCE	2 (bases 1 to 10242)		
AUTHORS	da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorfy, H., Faria, J.B., Ferreira, A.C.S., Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira, J.R.H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezra, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-NOV-2001) Departamento de Bioquimica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil		
FEATURES			
source	Location/Qualifiers		
	1. .10242		
	/organism="Xanthomonas axonopodis pv. citri str. 306"		
	/strain="306"		
	/db_xref="taxon:190486"		
gene	/note="pathovar: citri"		
	95. .286		
	/gene="XAC1412"		
CDS	95. .286		
	/gene="XAC1412"		
	/note="putative; ORF located using Glimmer/Genemark"		
	/codon_start=1		
	/transl_table=11		

Qy	651	GGCCTCGAGTCATCGAGAGCGCCGAATTCCTCGGCTTCAGAGCCGACCGGATCGAGGC	710
Ds	648	GGGCTGGAGTGATCGAGGCGCACCTGCTACGACGTGCGCTTCAGCGGATCGAGGT	707
Qy	711	GGTGGTCATCGCAATCCATCGCCATCGATCGATGGTGGCTTCGCGACGGGCGCTGAT	770
Ds	708	GGTGGTCATCGCGAGTCGGTTCATTCGATGGTGGAAATTCGTGACGGTTCGACGAT	767
Qy	771	GGCCCATCTCGGCCCGCCGACATCGCCACCGCATCGGATTCGCGTGAACCTGGCCGGG	830
Ds	768	GGCCAGCGCCAGCCCGCGGACATCGCATCGCATCGCGCTGGCCCTCGGCTGGCCGGA	827
Qy	831	TCGCGGCGAGTGCCCTCGCCCGGATCGAGCTCGACATCGACATTCGCGACCTCACTTCCA	890
Ds	828	CCGGTCCCGGACGCGCCCGCCGCTCGACATCGACATTCGCGACCTCACTTCCA	887
Qy	891	GAAGCCCTGACGAGAACGCTTTCGCGGCCCTGAGGCTTCGCGGAGACGCTCATCGGCGCG	950
Ds	888	CCGCTCGACACGAGCGGTTCGCGGCCGCTGAGCTTCGCGGAGGCTCGGCTCGGCTCGG	947
Qy	951	CGGCTCTCGGCGCCGCTTCACGCGGCGCAAGGATCGCGCTCGATCATTCATTCGCG	1010
Ds	948	CGGACCGCCCGCGCTCTCAATCGCGCAAGGAGGATGCTGCGACGCTTTCCTGAA	1007
Qy	1011	CGGACGATCGGTTTCTGACATCGCGCGGCTTCGCGGAGACGCTCGCGG	1063
Ds	1008	GGCGGCACTGCTTCACCGGATCGTGGACATCTGCGGAGGCTGCGCGG	1060
RESULT 15			
AE002375/c			
LOCUS	AE002375	13538 bp	DNA linear BCT 25-MAY-2000
DEFINITION	Neisseria meningitidis serogroup B strain MC58 section 17 of 206 of the complete genome.		
ACCESSION	AE002375	AE002098	
VERSION	AE002375.1	GI:725394	
KEYWORDS	Neisseria meningitidis MC58.		
SOURCE	Neisseria meningitidis MC58		
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
REFERENCE	1 (bases 1 to 13538)		
AUTHORS	Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clitstone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V., Pizzi, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.		
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58		
JOURNAL	Science 287 (5459), 1809-1815 (2000)		
MEDLINE	20175755		
PUBMED	10710307		
REFERENCE	2 (bases 1 to 13538)		
AUTHORS	Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clitstone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V., Pizzi, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712		
FEATURES	Medical Center Dr, Rockville, MD 20850, USA		
source	Location/Qualifiers		
	1. .13538		
	/organism="Neisseria meningitidis MC58"		
	/strain="MC58"		

/db_xref="taxon:125586"	
/note="serogroup: B"	
/complement(121..1377)	
/gene="NMB0176"	
/complement(121..1377)	
/gene="NMB0176"	
/note="similar to GB:102948 SP:P29011 PID:145703 CB:000096 PID:165191 percent identity: 76.98; identified by sequence similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="D-amino acid dehydrogenase, small subunit"	
/protein_id="AAF40633.1"	
/db_xref="GI:7225395"	
/translation="MKVLVAGVAGVSSAWYLAEAGHEVTVIDRAEGVAMETSFANA GOLSGYTTPWAAPGIPKALWLFKSHPLLPDPGSLYQIEMLMQMLQNCFAAHYQ TKERVRJSEYREMERFEQATGMFEGKGTQIQRQTEKVEAAKODIAVLERY GVPYRLKPECAEFAPALARTAKIAGLHLPADATGCRLEFENLYKLCQKGVRF HFQNLISRIDHGLRIKTVETETGFEADAVCALGCFSTKVLQAQDLNLPITPVKGI SUTLPVNSDGPVSTVDESIVKAITREFDNRIRVGGMALSGEYIKLPEKRETLAL VYNDLPPEGGLDQASQFWSGLRPMTPDSTPLIGTRFDNLFNTGHTGLGWTSLGSA KLTADIVSGKDEIRSDLSLSRYQA"	
/complement(1456..2847)	
/gene="NMB0177"	
/complement(1456..2847)	
/gene="NMB0177"	
/note="similar to GB:AE000511 PID:2314080 percent identity: 69.11; identified by sequence similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="sodium/alanine symporter, putative"	
/protein_id="AAF40634.1"	
/db_xref="GI:7225396"	
/translation="MOVFLDNPKAFETVSGVWVGLMLLVGTGILLTLVLLKGLQF TMLGYALKOAFMPKPKKSGEGEDISHAALMTALSTAGTGNAGVATVVTGGP GAVFNNWMTAIFGMATKYEGVLAIVRYVNNKSGMSGPMYIEKGLKNWKMVA FALFGTFASGIGSSVQSSVAQAVTSGIEPAYTGITLVITVAIVVGLGIGIKANA ASFIVPAMAVFVYVGLSIAINSALMPAVKLIFSDAFSAQAVAGAIIVRYGVA GVFNSNEAGMSAPIAAAKTDHPVQALVMTGFTLDTIVVCSITGIVLVMLLGA GEFVAPVEVSGAALTIVTFQKMLPGIGWITVIGLIFAYSTILGWCYGEKCAVVF GEKFAGLYRVGVSSVYMLGTVLSLDVLWLASDTFNGLMALPNLIALLLMAKIVNETR DFKQKITNGELPH"	
/complement(3204..3980)	
/gene="NMB0178"	
/complement(3204..3980)	
/gene="NMB0178"	
/note="similar to PID:1718489 percent identity: 100.00; identified by sequence similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase"	
/protein_id="AAF40635.1"	
/db_xref="GI:7225397"	
/translation="WTLIHPATAVIDPKAELDSGVKGVATVIGPNVQIGANTEIGPHA VINGHTSIGENNRIFQASGEIPEQDKYRDEPTKLIIGNGNTIRFETFTNLTGTGI GETRIDGDNWIMAYCHLADCVIGNHTIFANNASLAGHTIGDYVVLGYTLVQFCR IGDTAMTAAAGVKDVPYFNASGYRAEPAGLNSGEMRRNGFTAQISAVKDVYKTL YHRGIPFEAKADILRRATQAEFLAVFRDFAQSARGIIR"	
/complement(4067..4516)	
/gene="NMB0179"	
/complement(4067..4516)	
/gene="NMB0179"	
/note="similar to PID:1718488 percent identity: 100.00; identified by sequence similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="(3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase"	
/protein_id="AAF40636.1"	
/db_xref="GI:7225398"	
/translation="MDVQPIEAKDIQKLIIPHYRPLQDRITAFEPKMTLTAKNVT	

```
INERFQGHFDPDLPVPCVGLLIEAMAQCGTLAILSEGGKRENEFFFFAGIDEARFRK
QVDPDQDLVFEVLLTSRRGIGKNAKVDGQVAVEAIMCAKRVV"
/translation(4550. .5596)
/gene="NMB0180"
CDS
complement(4550. .5596)
/gene="NMB0180"
note="similar to PID:1718487 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="UDP-3-O-(3-hydroxymyristoyl)-glucosamine
N-acyltransferase"
/protein_id="AAF40637.1"
/db_xref="GI:7225399"
/translation="MIPATYITLSQITARLGGWGEDISVTAVRPLADAAEHISFLA
HPKYAEVHDSAGAVIVSAKAADGEGFRNLIVADDPYLYFAKVARLFPSPVVKARGGI
NPTAVPEGAVPTSCIEIGANVYIGANTLGEGRILANAVVQHDCKLGDSEVLHPNA
VYVYCTILGRREVHSGAVIGADGFLAFADDSWFKIPQTGAVTLGDDVEIGSNTNID
RGMSDTYVGNKIDNQVQIGHCKKIGSHVIAAKTGISGVSVTIGSYCIIGGGVGV
GHIEADKTTIGGGTSTVHSITGESKHLGIFIPMSTHKEWARNNAVYIHRUSEMNRKL
TLEQLSDAGQDSK"
complement(5629. .6129)
/gene="NMB0181"
CDS
complement(5629. .6129)
/gene="NMB0181"
note="similar to GB:M34854 SP:P31519 PID:155479 percent
identity: 46.88; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein OmpH, putative"
/protein_id="AAF40638.1"
/db_xref="GI:7225400"
/translation="MTRITRAFAALIGLCCTAGAHADTFQKIGFINTERIYLESKOA
RKIOKTLDESESAQDELQKREGLDLEROLAEGKLRNAKKAQAEKRWGLVAFRK
KQAFEEYDNRNEEFASLQNNRVIVKAKQEGDVLQNVYIYNTQIDVDSVI
KEMNAR"
complement(6195. .8588)
/gene="NMB0182"
CDS
complement(6195. .8588)
/gene="NMB0182"
note="similar to GP:2460281 percent identity: 99.87;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein Omp85"
/protein_id="AAF40639.1"
/db_xref="GI:7225401"
/translation="MKLKQIASALMLGISPLALADFTIQDIRVEGLQRTPEPSTVFN
LPVKVDYNDTHGSAIISKSYATGFFDDVRVETADGOLLITVIERPTIGSLNITGAK
MLQNDATIKLESFGLAOSQYFNQATLNQAVAGLKEEYLGKRLNIQITPKVTKLARN
RVDITITDEGSAKITDIEFEGNOVYSDRLMRQMSLTGGIWTWLTFRSNQFNEQKF
AODMEKVTDFYONNGYDFRILTDIOTNEBKTKOTIKITVHEGGRFWRGVSTEGDT
NEVPAKELEKLLMKPCGWYERQOMTAVLGEIONRMSGAVAYSEISVQPLPNAETKT
VDFVLTEPGRKIYVNEHITGNKTRDEVRVRELROMESAPYDTSKLQSKERVEILL
GFVNFQDFPLAGTDPKDLNLSLTERSTGSLDSAGWQDGTGLVNSAGVSQDNLF
GTGSAALRASRSTYLLNGSLSDPYFTADVGLSDYGVYKAFDPKRAKSTSIKQYKT
ITAGAGIRMSVPVETIDRYNFGSLVAEHLTVNTYKAPHYADFKIKYKTDGTGDSFK
FWLYKGTWGRNKTDSALWPTFRGLYGVNAEIALPGSKLOYYSATHNOTWFFPLSKT
TFLMGGEVLAGGYGRTKEIPFENEGYGGSLGSVRGESLTLPKPYVDEGEKISYC
GNKANVSAELLFPMPGAKDARTVRLSLFADAGSVWDGKTYDDNSSASTGVRQNIY
AGNTHKSTFTNELRYISAGAVTWLSPLGPMKFSYAYPLKKPEDEIQRFPOLGTTF"
complement(8645. .9985)
/gene="NMB0183"
CDS
complement(8645. .9985)
/gene="NMB0183"
note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40640.1"
/db_xref="GI:7225402"
```

```
/translation="MHTLLAFIFAILILVLSHERGHIIVARLCGVKVVYRFSVGGFKPF
FTRRGDTEWCLAPIPLGGLYKVMVDREGEVSEADLYAFDQKHPKARIIVAAAGPLT
NLALAVLILYGLSFSGVTLGPRYVGTVPDPTIARAGFQSGDKIQSVNGTIPVADMGSA
QTEIVLNLEAGVAVGVTASGATVRTIDAAGTPEAGKIAKNOCYIGLMPFKITTTVA
GGVEKGSPEAKGALPKPDRLLTAADGKPIASQEWANLTROSPGKITLNYBRAGOTHT
ADIRPDVTEQSDHTLILIGRVGLRQPDRADQAIIRKSRPSVVRFRMGWEKTVSHWT
TLKFFGKLISNASVSHISGLPITTIADIAGSAELGLOSTLEFLALVLSGLVNLPLV
PVLDDGHLIVFYTAEMIRGKPLIGERTVQIGLRFGLALMLMMAVAFNDVTRLGG"
complement(10020. .11204)
/gene="NMB0184"
CDS
complement(10020. .11204)
/gene="NMB0184"
Query Match 20.2%; Score 238.2; DB 1; Length 13538;
Best Local Similarity 53.8%; Pred. No. 2.1e-18;
Matches 598; Conservative 0; Mismatches 483; Indels 30; Gaps 4;
QY 9 CCGTGTGATCTTTGGGGCCACCGGCTCCATCGGCGAATCCACCTTCGACCTGTCATGCG 68
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11190 CCTGACCATTATTAGCAGTACCTACCGGAGCATAGGCGAAGCAGCTGAGCGTTCTCCC- 11132
QY 69 GAAGGGGGGGCCGAGGCGCTCCGCACCGCTCTGACCGGGCGGCGCAACATCCGGCG 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11131 -----GCCACCCCGAAAAATTCGCGTATTGCGCTGGCAGGCATAACAGGTCGAGAA 11077
QY 129 ACTGCGCGAAATGGCGCTGCGCTGAAGCGGAGCTTGCGCTCACCGGCATGAGGACTG 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11076 ATTGGCGGCTCAATGTCAAACGTTCCACCCGAATATGCGCTGTCGCGATCGGAACA 11017
...QY 189 CCGTCCCGCGCTGCGGAGGCGCGT-----GCCGGGAGCGGACGAGGTGCGGGCGG 242
Db 11016 CCGCCCGCGGCTTGAAGCGCTTGAACGCGGACGCGGAGCTCAAGCTTTTACACGG 10957
QY 243 GCGCAGGCGCATCGCGGAGGCGCGC---GACCGCGCGCGGCTGAGCAGCTGCGGCCAT 299
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10956 CCGCAGGCGATTGGTTGACGTTGCTCTGCCACGAAGTCACGGTGTCTATGTGCCCAT 10897
QY 300 CCGTGGCGCGCGCGGCGCTCGTCCCGGAATCGCGCGCTGAAAGCAGCGGCGCACGCTGCG 359
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10896 CGTCGGCGCGTGGGCGTCCCTCGCGCTCGACGCGGCGCAAAAGGCAAAACCATTTA 10837
QY 360 GCTCGCCACACAGAAAGCCTCGTACGCGGCGCAACTCTTGATCGGCGGCGCCAGCA 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10836 TCTGCAAAAGAAAGACGCTGGTGTTCGCGCGCGGCTTTATGGAACCGCGCGTGC 10777
QY 420 GAACGCGCGCACGATCTCGCGGTGGACAGCAGCAGCTCCGCGCTTTTCAGCGCGTGGC 479
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10776 AAACGCGCGCGCAGTCTCGCGCTCGACAGCAACACACGCGGTTTCCAGTTTGGC 10717
QY 480 GGGCAGGACACGG-----CCTGCGTGGAGCGCGTCTCATCATCAGCGCGTC 524
Db 10716 GCGCGATTACGCGCGCGCTCTGAACGCAACACGCGCATCGCTTATCTGACCGCTTC 10657
QY 525 CCGCGCGCGCTTCGCGACTGAGCGCTCGAGCGCATCCGCGCTGCACCGTGGCGGAGCG 584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10656 CCGCGCGCGCTTCTGACCGCGGATTAACACGTTTCGACCGCATTTACGCCGCCCAAGC 10597
QY 585 GCAGGCGCATCCCACTGTGCTCATGGCGGCGGATCTCCATCGACAGCGCGCTCGATGTT 644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10596 GGTCAACACCCCAATTGGCGTATGGGACGCAAAATCTCCGTCGATTCGCCACCATGAT 10537
QY 645 CAACAGGCGCTCGAGCTGATCGAGACGCGGAAATTTCTTGGCTTCGAGCGCGGACCGGAT 704
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10536 GAACAAAGTTTGGAGCTGATTGAAGCGCATTTGGCTGTTTCAACTGTCCCGCGGACAACT 10477
QY 705 CGAGCGGTGCTCCATCCCAATCCATCTCCATGCGATGGTGGGCTTCTGCGACGGGG 764
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10476 CCAAGTCGTCATCCATCCGCAATCCGTCATACAGCATGGTGGCGCTACCGGACGGCTC 10417
QY 765 CCGTATGGCGCCATCGGCGCGCGGACATGCGCCACGCGCATCGGATTCGCGCTGAACATG 824
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10416 CGTGTGGCGCAATGGGCAATCCCGATATGCAAGCGCCCATCGCTATTGTTGGGTTT 10357
```


Result	Query				Description	
No.	Score	Length	ID			
C	1	214.6	18.2	4403765	US-09-103-840A-2	Sequence 2, Appl
	2	126.8	10.7	1197	US-09-491-362-6	Sequence 6, Appl
	3	126.8	10.7	1197	US-09-874-562-5	Sequence 6, Appl
	4	125.2	10.6	1732	US-09-449-335-5	Sequence 5, Appl
	5	125.4	10.5	1434	US-09-480-921B-28	Sequence 28, Appl
	6	123.6	10.5	1732	US-09-449-335-1	Sequence 1, Appl
	7	122	10.3	1434	US-09-480-921B-9	Sequence 9, Appl
	8	99.2	8.4	1759	US-09-491-362-1	Sequence 1, Appl
	9	99.2	8.4	1759	US-09-874-562-1	Sequence 1, Appl
	10	74.2	6.3	8438	US-07-945-283-1	Sequence 1, Appl
	11	72.2	6.1	2809	US-09-000-016-1	Sequence 1, Appl
	12	72.2	6.1	2809	US-09-514-340-1	Sequence 1, Appl
	13	71.6	6.1	5173	US-08-242-677-1	Sequence 1, Appl
	14	69.4	5.9	20235	US-07-642-734C-3	Sequence 3, Appl
	15	69.4	5.9	20235	US-08-439-009A-3	Sequence 3, Appl
	16	66.6	5.6	2539	US-09-000-016-3	Sequence 3, Appl
	17	66.6	5.6	2539	US-09-514-340-3	Sequence 3, Appl
	18	66.6	5.6	11220	US-09-105-537-32	Sequence 32, Appl
	19	66.6	5.6	36778	US-09-105-537-5	Sequence 5, Appl
	20	66.6	5.6	38506	US-09-320-878-19	Sequence 19, Appl
	21	65.6	5.5	1288	US-08-440-856A-9	Sequence 9, Appl
	22	64.6	5.5	44377	US-08-804-227C-7	Sequence 7, Appl
	23	64.6	5.5	44377	US-08-804-198-1	Sequence 1, Appl
	24	64.6	5.4	1187	US-08-440-856A-2	Sequence 2, Appl
	25	63.2	5.3	1182	US-09-385-028-19	Sequence 19, Appl
	26	63.2	5.3	11604	US-09-385-028-13	Sequence 13, Appl
	27	63.2	5.3	15079	US-09-385-028-1	Sequence 1, Appl

; NAME/KEY: CDS
; LOCATION: (1)...(1197)
US-09-874-562-6

Query Match 10.78; Score 126.8; DB 4; Length 1197;
Best Local Similarity 49.08; Pred. No. 1.9e-15;
Matches 375; Conservative 0; Mismatches 382; Indels 9; Gaps 1;

```
Oy 293 CGGCCATCGTGGGCGCGGCGCTCGTCCCGGAATCGGCGCTGAAGCAGCGCCGCA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 CCGGAATAGTAGTGTGCGGCGACTAAAGCCTACGGTGTGCAATGAAGCAGGAAGG 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 353 CGTGGCGCTGCCAACAGAAAGCCCTGTCAGCGGAGGCACTCTGTATGCGGAGG 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 ACATTGCTTTGCAACAAAGAGACATTAATCGCAGGTGGTCTTTCGTGCTCCGCTG 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 413 CCCAGGAGAACGGCCACGATCTGCGGTGGACAGCGACCTCCGCGGTCTTTCAGG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 CCAACAACATAAATGAAGATTTCTCGGCGAGATTGAGACATCTGCCATATTTCAAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 473 CGTGGGCGGAGGACACGCGCTCGTTCGAGCGCTCATCATCACGCGCTCCGCGGCG 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 GTATTCAAGGTTTGCTGAAGCGCTCTGCGCAAGATAATCTTGACTGCATCTGTGGAG 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 533 GCTTCCGCGACTGGAGCCTCAGCGCATCCCGCGCTGACCGTGGCGGAGCGGCGCC 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 CTTTTAGGATTTGGCTGTGAAAAGCTTAAAGGAAGTTAAAGTAGCGGATCGTTGAAG 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 593 ATCCCACTGTGTCATGGCGGCGGATCTCCATCGACAGCGCTCGATGTTCAACAAG 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 ATCCAACTGGAATGGAAGAAATCACTGTGGACTCTGCTACGCTTTTCAACAAG 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 653 CGCTCGAGCTGATCGAGACGCGCAATTTCTTGGCTTCGAGCGGACCGGATCGAGCGG 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 GTCCTGAGTCTATTGAAGCGCATTTGTTTGGAGCTGAGTATCAGCATATAGAGATTG 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 713 TCGTCCATCCCATCCATCTCCATCGATGATGGTGGGCTTCTGGAAGGGGCGCTGATGG 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 TCATTCTCCGCAAGTATACATATCCATGATTGAAACACAGGATTCATCTGTGCTTG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 773 CCCATCTCGGCGCGCGGACATGGCCACGCGATCGGATTCGCGTGAACCTGGCGGCTC 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 779 CTCATTTGGTGTGCTGATGCGTTTACCGATTCTACACCATGTGATGGCCGATA 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 833 GCGGC-----GAGGTGCGGTCGCGGATCGACTCGACAGATTCGAGCGCTCA 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 839 GAGTTCCTTGTCTGAAGTAATTTGGCCAAAGACTTGACCTTTGCAAGCTCGGTTCAATGA 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 884 CTTCCAGAGCCTGACAGGACGCTTTCGCGCCTGAGGCTTCGCGGAGACGTCATGG 943
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 899 CTTTCAAGAAACAGACAATGTGAATACCCATCCATGGATCTTGTGCTTGAC 958
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 944 CGGCGCGGCGCTGTCGGGCGCGCTTCAACGCGGCGCAAGAGATCGCGCTCGATCAT 1003
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 959 GAGCTGGAGGCACATGACTGGAGTTCTCAGCGCGCCCAATGAGAAAGCTGTTGAATGT 1018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1004 TCATCGCGGAGCATCGGTTTCTGAGACATGGCGGCGGTGCTGA 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1019 TCATTGATGAAGATAAGACTTTGGATATCTTCAAGGTTGTGA 1064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4

US-09-449-335-5
; Sequence 5, Application US/09449335
; Patent No. 6303365
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Kloti, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and

; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/449,335
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: DE 199 35 967.9
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: new coding sequence
US-09-449-335-5

Query Match 10.68; Score 125.2; DB 4; Length 1732;
Best Local Similarity 48.88; Pred. No. 3.9e-15;
Matches 374; Conservative 0; Mismatches 383; Indels 9; Gaps 1;

```
Oy 293 CGGCCATCGTGGGCGCGGCGCTCGTCCCGGAATCGGCGCTGAAGCAGCGCCGCA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 CCGGAATAGTAGTGTGCGGCGACTAAAGCCTACGGTGTGCAATGAAGCAGGAAGG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 353 CGTGGCGCTGCCAACAGAAAGCCCTCGTTCAGCGGAGGCACTCTGTATGCGGAGG 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 ACATTGCTTTGCAACAAAGAGACATTAATCGCAGGTGGTCTTTCGTGCTCCGCTG 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 413 CCCAGGAGAACGGCCACGATCTCGCGGTGGACAGCGACCTCCGCGGTCTTTCAGG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 CCAACAACATAAATGAAGATTTCTTCGCGGAGATTGAGACATTTCTGCCATATTTCAAT 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 473 CGCTGGCGGCGGAGGACACGCGCTCGCTCGAGCGCTCATCATCACGCGCTCCGCGGCG 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 GTATTCAAGGTTTGCTGAAGCGCTCTCGCAAGATAATCTTGACTGCATCTGTGGAG 789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 533 GCTTCCGCGACTGGAGCCTCAGCGCATCCCGCGCTGACCGTGGCGGAGCGGAGGCC 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 790 CTTTTAGGATTTGGCTGTGAAAAGCTTAAAGTAAGTAGCGGATCGGTTGAAGC 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 593 ATCCCACTGTGTCATGGCGGCGGATCTCCATCGACAGCGCTCGATGTTCAACAAG 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 850 ATCCAACTGGAATGGAAGAAATCACTGTGAGCTCTGCTAGCTTTTCAACAAG 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 653 CGCTCGAGCTGATCGAGACGCGCAATTTCTCGGCTTCGAGCGGACCGGATCGAGCGG 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 910 GTCCTGAGTCTATTGAAGCGCATTTGTTTGGAGCTGAGTATGAGATATAGAGATTG 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 713 TCGTCCATCGGCAATCCATGTCATGCGATGCGGCTTCTGCGAGCGGCGGCTGATGG 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 970 TCATTCTCCGCAAGTATACATATCCATGATTGAAACACAGGATTCATCTGTGCTTG 1029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 773 CCCATCTCGGCGCGCGGACATGGCCACGCGCATCGGATTCGCGCTGAACCTCGAGCGCTCA 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1030 CTCATTTGGTGTGCTGATATGCGTTTACCAGTTCTCTACACCATGTGATGCGCCGATA 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 833 GCGGC-----GAGGTGCGGCTCGCGCGGATCGACTCGACAGATTCGAGCGCTCA 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1090 GAGTTCCTTGTCTGAAGTAATTTGGCCAAAGACTTGACCTTTTCAACAAGCTGTTCAATGA 1149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 884 CTTCCAGAGCCTGACAGGAAACGCTTTCGCGGCGCTGAGGCTTGGCGAGACGTCATGG 943
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1150 CTTTCAAGAAACAGACAATGTGAATACCCATCCATGAGTCTTGTGCTTGTGCGGAG 1209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 944 CGGCGCGGCGCTGTCGGGCGCGCTTCAACGCGGCGCAAGAGATCGCGCTCGATCAT 1003
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1210 GAGCTGGAGGCACATGACTGGAGTTCTCAGCGCGCCCAATGAGAAAGCTGTTGAATGT 1269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1004 TCATCGCGGAGCATCGGTTTCTGAGACATGGCGCGCGGTGCTGA 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1270 TCATTGATGAAGATAAGCTATTTGGATATCTTCAAGGTTGTGA 1315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 5
US-09-480-921B-28
; Sequence 28, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1434)
; OTHER INFORMATION: encodes SEQ ID NO:29
US-09-480-921B-28

Query Match 10.5%; Score 124.4; DB 4; Length 1434;
Best Local Similarity 48.6%; Pred. No. 5.4e-15;
Matches 372; Conservative 2; Mismatches 383; Indels 9; Gaps 1;

QY 293 CGGCCATCTGGGGCGCCGCGCTCTGTCGGGAATGCGGGCGCTGAAGCAGCGCGCA 352
Db 533 CCGGAATAGTAGTTGTGGCGGACTAGCGGTGTGCAATTGAAGCAAGG 592
QY 353 CGCTGGCGCTCGCAACAAGAAAGCCTCTGACGGCAGCACTCTCTATCGGACGG 412
Db 593 ACATTGCTCTGTGCAACAAGAGACATTAATCGCAGGTGTCTTTCGTCTTCCGCTG 652
QY 413 CCCAGGAGAACGGCCGACGCTCTGCGGTGGACAGCAGCACTCGCGGTCTTTCAGG 472
Db 653 CCAACAACATAATGTAAGATCTTCCGCGAGATTCAGAACATCTGCCATATTTTCA 712
QY 473 CGTGGCGGCGGAGGACAGCGCTCTGCGGTGGACAGCAGCACTCGCGGTCTTTCAGG 472
Db 713 GTATTCAAGTTTGCCTGGAAGCGCTCTGCGCAAGATAATCTTGACTGCACTTGTGGAG 772
QY 533 CGTTCCGCGACTGGAGCTCGAGCGCATCGCGCTGACCGTGGCGGAGCGCAGGCC 592
Db 773 CTTTGGGATTGGCTGTCGAAAGCTAAGGAAGTTAAGTAGCGGATGCTTGAAGC 832
QY 593 ATCCCAACTGGTCCATGGGCGCAGCGATCTCCATCGACAGCGCTCGATGTTTCAACAAG 652
Db 833 ATCCAAACTGGAACATGGGAAGAAATCACTGTGACTCTGTACGCTTTTCAACAAG 892
QY 653 CGCTCGAGCTGATCGAGACCGCGAATCTTTCGGCTTCGAGCGGACCGGATCGAGCGG 712
Db 893 GTCTTGAAGTCATGAAGCGCATATTTGTGGAGCTGAGTATGACGATATAGATTG 952
QY 713 TCGTCATCGCAATCCATCGTCCGATGTGGCTTCTCGCAGCGGGGCTCTATGG 772
Db 953 TCATTATCCGAAGTATCATATCATCTCCATGATTGAACACAGGATTCATCTGTGCTTG 1012
QY 773 CCATCTCGGCCCGCGCATCGCCACGCCATCGGATTCGGCTGAACTGGCGGGTTC 832
Db 1013 CTCAATTGGGTTGGCTGATATCGGTTTACCGATTCTCTACACCATGTCTATGSCCGATA 1072
QY 833 GCGG-----CGAGTGGCGCTCGCCGGATCGACCTCGCACAGATTCGAGCGCTCA 893
Db 1073 GAGTTCCTTGTCTGAAGTAACCTTGCCWAGACTTGACCTTTGCAAACTCGGTTTCATGA 1132
QY 884 CTTTCCAGAAAGCTGACGAGGAACGCTTTTCGGCGCTGAGGCTTTCGCGAGAGCTATGG 943

Db 1133 CTTTCAAGAAACCAGACAATGTGAATACCATCCATCGATCTTGTCTTATGCTGCTGAC 1192
QY 944 CGGCGCGCGCTGTGCGGGCGCCCTTCAACGGCGCAAGGAGATCGCGTCTCATCAT 1003
Db 1193 GAGCTGAGGCACAACTACTGGAGTTCTCAGCGCGCAATGAGAAAGCTGTGAAATGT 1252
QY 1004 TCATCGCGGACGATCGGGTTTCTTGACATGCGCGCGGTGGTGA 1049
Db 1253 TYATTGATGAAAGATAAGCTATTGATGATATCTTCAAGGTTGTGA 1298

RESULT 6
US-09-449-335-1
; Sequence 1, Application US/09449335
; Patent No. 6303365
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Kloti, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/449,335
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: DE 199 35 967.9
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: DNA encoding 1-deoxy-D-xylulose-5-phosphate
; OTHER INFORMATION: reductoisomerase
US-09-449-335-1

Query Match 10.5%; Score 123.6; DB 4; Length 1732;
Best Local Similarity 48.7%; Pred. No. 7.6e-15;
Matches 373; Conservative 0; Mismatches 384; Indels 9; Gaps 1;

QY 293 CGGCCATCTGGGGCGCCGCGCTCTGTCGGGAATGCGGGCGCTGAAGCAGCGCGCA 352
Db 550 CCGGAATAGTAGTTGTGCGGGACTAAGCCTACGGTTGCTGCAATTGAAGCAGGAAGG 609
QY 353 CGTGGCGCTCGCAACAAGAAAGCCTCTGACGGCAGGGCACTCTGATCGGACGG 412
Db 610 ACATTGCTCTTGGCAACAAGAGACATTAATCGCAGGTGTCTTTCGTCTTCCGCTG 669
QY 413 CCCAGGAGAACGGCCGACGATCTCGCGGTGGACAGCAGCACTCGCGGTCTTTCAGG 472
Db 670 CCAACAACATAATGTAAGATCTTTCGGCAGATTTCAGAACATCTGCAATATTTCA 729
QY 473 CGCTGGCGGAGGACAGCGCTCTGCGCGCTGAGCGCTCATCATCAGCGCTCGCGGGC 532
Db 730 GTATTCAAGTTTGCCTGAAGCGCTCTGCGGAAGATAATCTTCTGCTGCTGCTGGAG 789
QY 533 CGTTCGCGACTGAGCCCTCGAGCGCATCGCGCTGACCGCTGCGCGGCGAGCGCGCC 592
Db 790 CTTTGGGATTGGCTGTGCAAAAGCTAAGGAAGTTAAAGTAGCGGATGCTTGAAGC 849
QY 593 ATCCCAACTGGTCCATGGCGCAGCGGATCTCCATCGACAGCGCTCGATGTTTCAACAAG 652
Db 850 ATCCAAACTGGAACATGGGAAGAAATCACTGTGACTCTTTCAGCTTTTCAACAAG 909
QY 653 CGCTCGAGCTGATCGAGACGCGGAATCTTTCGGCTTCGAGCGCGGACCGGATCGAGCGG 712
Db 910 GTCTTGAAGTCATTGAAGCGCATATTTGTTGGAGCTGAGTATGAGATATGAGATTG 969

QY 713 TCGTCATCCGCAATCCATCGTCCATGCGATGGTGGGCTTCTGCGACGGGGCGCTGATGG 772
Db 970 TCATTTCATCCGCAAGTATCATACATTCATGATGAAACACAGGATTCATCTGCTTG 1029
QY 773 CCCATCTCGGCGCGCGACATGCGCCAGCCATCGGATTCGGCTGAACTGCGCGGGTC 832
Db 1030 CTCATTTGGTGGCGCTGATATGCGCTTTACCGATTCCTACACCATGTCATGCGCCGATA 1089
QY 833 GGGC-----GAGGTGCGCGCTGCGCCGATCGACCTCGCACAGATTGCGAGCCTCA 883
Db 1090 GAGTTCCTTGTCTGAAGTAACTTGCCCAAGACTTGACCTTTGCAAACTCGGTTCAATGA 1149
QY 884 CTTTCAGAAAGCTGACGAGGAAAGCTTTTCGGGCGCTGAGGCTGCGCGAGACGTCATGG 943
Db 1150 CTTTCAGAAAGCAAGCAATGTGAATACCCATCCATGATCTGCTTATGCTGCTGGAC 1209
QY 944 CGGCGCGCGCGCTGTCGGGCGCGCTTCAAGCGGCGCAAGGAGATCGGCTGATCATTT 1003
Db 1210 GAGCTGGAGGCAATGACTGGAGTTCTCAGCGCGCGCAATGAGAAAGCTGTTGAAATGT 1269
QY 1004 TCATCGCGCGAGCATCGGCTTCTGGACATGGCGCGCGTGGTGA 1049
Db 1270 TCATTGATGAAGATAAGCTATTGGATATCTTCAAGTTGTGA 1315

RESULT 7

US-09-480-921B-9
; Sequence 9, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
US-09-480-921B-9

Query Match 10.3%; Score 122; DB 4; Length 1434;
Best Local Similarity 48.6%; Pred. No. 1.5e-14;
Matches 372; Conservative 0; Mismatches 385; Indels 9; Gaps 1;

QY 293 CGGCCATCGTGGCGCGCGGCTCGTCCCGGAATGCGGCGCTGAAGACGCGCGCA 352
Db 533 CCGGAATAGTAGTGTGCGGCGACTGAAGCCTACGGTTGCTGCAATTGAAGCAGGAAAGG 592
QY 353 CGCTGCGCTCGCAACAGGAAGCCTCGTGACGCGCAGGCAACTCTGATCGGACGG 412
Db 593 ACATTGCTCTTGAACAAGAGACATAATCGCAGGTGGCTTTCGTGCTTCGCTTG 652
QY 413 CCAGGAGAACGCGCGCAGCATCTCGCGTGGACAGCAGCACTCCGCGGCTTTTCAGG 472
Db 653 CCAACAACATATGTAAGATTTCTCGGCAGATTTCAGAACATTCGCCATATTTTCAGT 712
QY 473 CGCTGGCGGCGAGACAGCGCTCGGTGAGCGCTCATCATCATCGGCGCTCGGCGGCG 532
Db 713 GTATTCAAGTTTGGCTGGAAGGCGCTCTCGGCAAGATAATCTTGACTGTCATCTGGTGGAG 772
QY 533 CGTTCCGCGACCTGAGCCCTCGAGCGCATCGGCGCTGACCGGTGGCGGAGGCGCGCC 592
Db 773 CTTTAGGATTTGGCTGTGCAAAAGCTAAAGAGTTAAAGTAGCGGATGCGTTGAAGC 832

QY 593 ATCCCAACTGGTCCATGGGCCAGCGATCTCCATCGACAGCGCTCGATGTTTCAACAAGG 652
Db 833 ATCCAACTGGAACATGGGAAAGAAATCACTGTGGACTCTGCTAGCTTTTCAACAAGG 892
QY 653 CGCTGAGCTGATGAGAGCGCGGAATTTCTCGGCTTCCAGCGGACCGGATCGAGGCGG 712
Db 893 GTCTTGAAGTCAATGAAGCGCATTTATTTGTTGGAGCTGAGTATGACGATATAGAGATTG 952
QY 713 TCGTCCATCCGCAATCCATCGTCCATGCGATGTTGGGCTTCTCGGACGGGCGCTGATGG 772
Db 953 TCATTTCATCTCAAGTATCATACATTCATGATGAAACACAGGATTCATCTGCTGCTTG 1012
QY 773 CCATCTCGGCGCGCGGACATGCGCCAGCCATCGGATTCGCGCTGAATCGCCGGGTC 832
Db 1013 CTCATTTGGTGGCGCTGATATGCGTTTACCGATTCCTACACCATGTCATGCGCCGATA 1072
QY 833 GCG-----GCGAGGTGCGCGCTGCGCGGATCGACCTCGCACAGATTGCGAGCCTCA 883
Db 1073 GAGTTCCTTGTCTGAAGTAACTTGGCCTAGACTTTCACCTTTGCAAACTCGGTTCAATGA 1132
QY 884 CTTTCAGAAAGCTGACGAGGAAAGCCTTTTCGGCGCTGAGGCTTGGCGGAGCGTCAATGG 943
Db 1133 CTTTCAGAAACACAGCAATGTGAATACCCATCCATGATCTGCTTATGCTGCTGGAC 1192
QY 944 CGGCGCGCGCGCTGTCGGCGCGCGCTTCAACGCGCGCAAGGAGATCGCGCTGATCATTT 1003
Db 1193 GAGCTGGAGGCAATGACTGGAGTTCTCAGCGCGCGCAATGAGAAAGCTGTTGAAATGT 1252
QY 1004 TCATCGCGCGAGCATCGGCTTCTGGAGATGCGGCGCGGCTGGTGA 1049
Db 1253 TTATTGATGAAGATAAGCTATTGGATATCTTCAAGGTTGTGA 1298

RESULT 8

US-09-491-362-1
; Sequence 1, Application US/09491362
; Patent No. 6281017
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: WSUR14977
; CURRENT APPLICATION NUMBER: US/09/491,362
; CURRENT FILING DATE: 2000-01-26
; EARLIER APPLICATION NUMBER: 60/118,349
; EARLIER FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(1496)
US-09-491-362-1

Query Match 8.4%; Score 99.2; DB 4; Length 1759;
Best Local Similarity 48.3%; Pred. No. 2.3e-10;
Matches 377; Conservative 0; Mismatches 388; Indels 15; Gaps 3;

QY 293 CGGCCATCGTGGCGCGCGGCTCGTCCCGGAATGCGGCGCTGAAGACGCGCGCA 352
Db 598 CGGAATTCGCGCTGCGAGTTTGAAGCCGACAGTGGCTGCCATAGAGCTGGAAGG 657
QY 353 CGCTGCGCTCGCAACAAGAAAGCCTCGTGACGCGGCAACTCTGATCGGACGG 412
Db 658 ACATTGCTTTGGCAATAAAGACACATAATCGCTGAGGCGCTTTGCTTCCTCTTG 717
QY 413 CCAGGAGAACGCGCGCCACGATCCTCGCGTGGACAGCAGCACTCGGCGGCTTTTCAGG 472

Db 718 CAAGAAGACACAACTCAAGATCTTCTCGCAGATCCGAACTCTGATATTTCACT 777
Qy 473 CGCTGGCGGAGACACGCGCTCGCTCGAGCGGCTCATCATCAGCGCTCGCGCGGC 532
Db 778 GTATCCAGGCTTGCCAGAGGTCTTTGAGCGGTATTAATTTGACTGCATCGGAGAG 837
Qy 533 CGTTCGCGACTGGAGCCTCGAGCGCATCCGCGCTCGACCGTGGCGAGCGCAGGCC 592
Db 838 CTTTCAGGGAATTTGCCGCTTGAGAAATGAAAGAGGTGAAAGTAGCAGATCTTTAAAGC 897
Qy 593 ATCCCACTGTCTCATGGCGCAGCGGATCTCCATCGACAGCGCCTCGATGTTTC---AACA 649
Db 898 ATTCCACTGGAATATGGGAAAAGAAATACAGTCCGACTTCTGCAACTCTTCTTTAACA 957
Qy 650 AGGCGCTCGAGCTGATCGAGACGCGCGAAATCTTCGCGCTTCGAGCGGACCGGATCGAGG 709
Db 958 AGGCGCTCGAAGTCATAAAGCTCACTATTTGTTGGGCGAGAATATGATGATTTGAGA 1017
Qy 710 CGGTGCTCCATCCGCAATCCATGCTCCATGCGATGCGGTGGGCTTCTGGCAGCGGCGCTGA 769
Db 1018 TTGTTATTCATTCCTCCATCCATCATCTACTCGATGGTCCGAGACACAGGATTCATCGGTGC 1077
Qy 770 TGGCCCATCTCGGCGCGCGCGCATGCGCCAGCGCATCGGATTCGCGTGAAGTGGCGG 829
Db 1078 TAGCTCAATAGGATGGCGCGCGATGCGTTGCGCTTATCTGTACACTTATCATGGCGAG 1137
Qy 830 -----GTGCGGCGGAGGTGCGCGCGCGCTTCGCGCTTCGAGCGGACCGGATTCGAGGCC 880
Db 1138 AGAGAGTCTACTGCTCCGAGATTACATGCGCTCGACTCGACCTCTGCAAGGTG---ATT 1194
Qy 881 TCACCTTCAGAGCTGACAGAGACGCTTTCCGCGCGCTGAGGCTTGGCGGAGAGCTCA 940
Db 1195 TACCATTCAAGAAGCGCGATACCGGTGAATACCGCTATGGATCTAGCCTATGCTGCTT 1254
Qy 941 TGGCGCGCGCGCTGCGCGCGCGCTTCGCGCGCGCGCAAGAGAGATCGCGCTCGATC 1000
Db 1255 GGAAGAGCGGAGCACCATGACCGGAGTCTGAGCGCAGCTAATGAGAAACAGTCGAAA 1314
Qy 1001 ATTTATCGCGGAGCATCGGTTTCTGGACATGCGCGCGGTGGTGGAGGAGAGCTCG 1060
Db 1315 TGTTCATCGAGAGAAATCGCTACCTCGACATTTTCAAGGCTGCTGGAGCTTACATGCG 1374

RESULT 9

US-09-874-562-1
; Sequence 1, Application US/09874562
; Patent No. 6420159
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; FILE REFERENCE: WSUR17549
; CURRENT APPLICATION NUMBER: US/09/874,562
; PRIOR FILING DATE: 2001-06-04
; PRIOR FILING DATE: 2000-01-26
; PRIOR FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(1496)
US-09-874-562-1

Query Match 8.4%; Score 99.2; DB 4; Length 1759;
Best Local Similarity 48.3%; Pred. No. 2.3e-10;
Matches 377; Conservative 0; Mismatches 388; Indels 15; Gaps 3;

Qy 293 CGGCATCTGCGCGCGCGGCTCTGTCCCGAAATCGCGGCGCTGAAGCACGCGCGCA 352
Db 598 CGGAATTTGCTGGCTGTGAGGTTTGAAGCCGACAGTGGCTGCCATAGAACTGGAAGG 657
Qy 353 CGTGGCGCTCGCCAAAGAAAGCCTCTGTGACGGCAGGCAACTCTCTGATGGGAGCG 412
Db 658 ACATTTGCTTTGGCCAAATAAGAGACACTAATCGCTGGAGGCGCTTTTGTCTCTCTCTG 717
Qy 413 CCCAGGAGAGCGCGCCACGATCTCTCGCGGTGGACAGGAGCAGCTCCGCGGTCTTTCAGG 472
Db 718 CAAGAAGACAACTCAAGATCTTCTCGACAGCTCCGAACTCTGCTATATTTCACT 777
Qy 473 CGCTGGCGGCGAGACACGCGCTGCGTCCGAGCGCTCATCACGCGCTCCGCGCGGC 532
Db 778 GTATCCAGGCTTCCAGAGGTCTTTGAGCGGTATAATTTTGTGCTCATCGGAGAG 837
Qy 533 CGTTCCGCGACTGGAGCTCGAGCGCATCCGCGCTGACCGTGGCGGAGCGGAGGCC 592
Db 838 CTTTCAGGGAATTTGCCGCTTGAGAAATGAAAGAGGTGAAAGTAGCAGATCTTTAAAGC 897
Qy 593 ATCCCACTGTCTCATGGCGCAGCGGATCTCCATCGACAGCGCCTCGATGTTTC---AACA 649
Db 898 ATTCCACTGGAATATGGGAAAAGAAATACAGTCCGACTTCTGCAACTCTTCTTTAACA 957
Qy 650 AGGCGCTCGAGCTGATCGAGACGCGCGAAATCTTCGCGCTTCGAGCGGACCGGATTCGAGG 709
Db 958 AGGCGCTCGAAGTCATAAAGCTCACTATTTGTTGGGCGAGAATATGATGATTTGAGA 1017
Qy 710 CGGTGCTCCATCCGCAATCCATGCTCCATGCGATGCGGTGGGCTTCTGGCAGCGGCGCTGA 769
Db 1018 TTGTTATTCATTCCTCCATCCATCATCTACTCGATGGTCCGAGACACAGGATTCATCGGTGC 1077
Qy 770 TGGCGCGCGCGCTGCGCGCGCGCTTCGCGCGCGCTTCGCGCTGAGCTGAGCTGCGCGG 829
Db 1078 TAGCTCAATAGGATGGCGCGCATGCGTTGCGCTTATCTGTACACTTATCATGGCGAG 1137
Qy 830 -----GTGCGGCGGAGGTGCGCGCGCGCTTCGCGCGCGCTCGACCTCGACAGATTCGAGGCC 880
Db 1138 AGAGAGTCTACTGCTCGGAGATTACATGCGCTCGACTCGACCTCTGCAAGGTG---ATT 1194
Qy 881 TCACCTTCAGAGAGCTGACAGAGACGCTTTCCGCGCGCTGAGGCTTGGCGGAGAGCTCA 940
Db 1195 TACCATTCAAGAAGCGCGGATACCGGTGAATACCGCTATGGATCTAGCCTATGCTGCTT 1254
Qy 941 TGGCGCGCGCGCTGTCGCGCGCGCGCTTCGCGCGCGCGCAAGAGAGATCGCGCTCGATC 1000
Db 1255 GGAAGAGCGGAGCACCATGACCGGAGTCTGAGCGCAGCTAATGAGAAACAGTCGAAA 1314
Qy 1001 ATTTATCGCGGAGCATCGGTTTCTGGACATGCGCGCGGTGGTGGAGGAGAGCTCG 1060
Db 1315 TGTTCATCGAGAGAAATCGCTACCTCGACATTTTCAAGGCTGCTGGAGCTTACATGCG 1374

RESULT 10

US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The Ep0 and LfT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8438 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 622..6495
FEATURE:
NAME/KEY: variation
LOCATION: replace(1099, "g")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1267, "t")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1381, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1566, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(7010, "g")
US-07-945-283-1

Query Match 6.3%; Score 74.2; DB 1; Length 8438;
Best Local Similarity 45.1%; Pred. No. 9,4e-06;
Matches 320; Conservative 0; Mismatches 383; Indels 6; Gaps 1;

QY 10 CTGTGATCTTTTGGGGCCACCGGCTCCATTCGGCGAATCCACCTTCGACCTCGTCATGCGG 69
DB 4977 CTGCGGCACATCGGGGAGCCCATGGCGCCAGGACCGCTCTGGGCGCTCGCCACGCG 4918

QY 70 AAGGCGGGCGGAGCGTTCGGACCGTCTGACCGGGGGCGCAACATCGGCGCA 129
DB 4917 GTGAGCGCGGTGGCCATGAGCGCGCGCTACGATCGCACCCAGAACCTTCATCTCGAG 4858

QY 130 CTGGCCGAATCTGGCGTGAAGCGGAGCTTGCGCTCACCGCGCATGAGGACTGC 189
DB 4857 AGCCTCGCGGGCTACGCGACATGGCTTACCGGGCGCGCGGACCCCGCGCC 4798

QY 190 CTGCGCGCGTTCGGGAGCGCTGGCGGAGCGGACCGAGTTCGGGGGGGGCGGAG 249
DB 4797 GGGGAGGAGCGCTGAGGCGCTTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGGAG 4738

QY 250 GCCATCGCGGAGCGCGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 309
DB 4737 CCGCGCGCGTTCGCGGAGCTGGCGGAGCTTCGCTCTGCGGCGGCGGCGGCGGCGG 4678

QY 310 GCGGGCGCTCTGCTCGCGGAGTGGGGGCTGAAGCACGCGCGGCGGCGGCGGCGGCGG 369
DB 4677 GAGCGCGTCTGCTCGCGGCTCGCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCTC 4618

QY 370 AAGGAAAGCTGTGACGGCAGGCGCACTCTGTATGCGGACGGCGGCGGCGGCGGCGG 429
```

```
DB 4617 GGGCGCGCTGCTCGAGGAGGTGACCGCGCGCTGCTCGCGTCCGACGCGATCCCC 4558
QY 430 ACGATCTCTCGCGTGGACAGCGAGCACTCCGCGGTCTTTTTCAGGCGCTGCGGCGGAGGAC 489
DB 4557 GGGCGCGCGCGCGGAGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4498
QY 490 ACGGCTGCTGCGTGGAGCGGCTCATCATCAGGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGG 549
DB 4497 GTGGCGCGCTGCTGCGCTACAGCGTGGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4438
QY 550 CTCGAGCGGATCCGCGCGCTGACCGTGGCGGCGGAGGCGGAGCGGCGGCGGCGGCGGCGGCGG 609
DB 4437 TACGCGCGGCGCTCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4378
QY 610 GCGCAGCGGATCTCCATCGACAGCGCTCGATGTTCAACAAGCGGCTCGAGCTGATCGAG 669
DB 4377 GCGCGCGCGGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4324
QY 670 ACGCGCGAATCTTCGCGCTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 718
DB 4323 CAGCGCGGCTGCTGCTCCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4275
```

RESULT 11

US-09-000-016-1

; Sequence 1, Application US/090000016

; Patent No. 6143541

; GENERAL INFORMATION:

; APPLICANT: AKIRA ARISAWA et al.

; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC

; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DE

; TITLE OF INVENTION: ITS EXPRESSION PRODUCT

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., #800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/000,016

; FILING DATE: January 30, 1998

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warren M. Cheek, Jr.

; REGISTRATION NUMBER: 33,367.

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-721-8200

; TELEFAX: 202-721-8250

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2809 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; ORIGINAL SOURCE:

; ORGANISM: Streptomyces viridosporus

; STRAIN: A-914

; ORIGINAL SOURCE:

; ORGANISM: Streptomyces antibioticus


```
Qy 447 CAGCAGCACTCCGGGGTCTTTTACGGCGCTGGCGGGGAGGACACAGCGGCTCGTTCGAGCG 506
Db 1966 CGCCACCGACCAAGGCAAGGGCGCCCGCGGCTTCTTACGCTGGCGCGCACACCA 2025
Qy 507 CGTATCATACAGCGCTCCGGCGGCGCTTCCGGGACTGGAGCCTCGAGCGCATCCGCGC 566
Db 2026 GGTACCTCCCGGGCGGCGAGCGCTTCCGTGACATACCGCCGACACCGGCTCGG 2085
Qy 567 CTGACCGCTCGGAGGCGGAGGCGCCATCCCAACTGTGTCCATGGCCACAGCGATCTCCAT 626
Db 2086 CGGCACGTGACGCGCGTACTCGCGCTAGCTGGTCCCGACGGCGGCGGCGACAGCGT 2145
Qy 627 CGACAGCGCTCGATGTTTCAACAGGCGCTCGAGCTGATGAGAGCGCGCGAATTTTCGG 686
Db 2146 CCGACGCGCGCGGCTGACGCGGAGTGTGAGTGTGACGCTGACCGTCCGCGCACAT 2205
Qy 687 CTTGAGCGCGACCGATCGAGGCGGTCTCATCCCGCAATCCATCGTCCATGCGATGCT 746
Db 2206 CGCGCGGAGCGCAAGCCACGACCGCAACACTCACCAGCTGATCGGCTACGCGGGCCT 2265
Qy 747 GGGCTTCTGCGACGCGGCGCTGATGGCCCATCTCGGCCCGCCGACATGCGCCACGCGCAT 806
Db 2266 GGGCTCGCGCGGTTACGCGCGCCCGCCGACCGACACCGCCACCTTGCCTTGCCCAA 2325
Qy 807 CGGATTCGCGTGAACCTGCGGCGGTCGCGGAGTGCCTCGCCCGGATCGACCTCGC 866
Db 2326 GGGCACCTTACCTGTGTGACTCTGTGATCGCCAAAGACTTCGGAGCTCAAGGGCGGAT 2385
Qy 867 ACAGATTCGAGCTCACCCTCCAGAGCTGACGAGAAAGCTTTCGGGCGCTGAGCT 926
Db 2386 CGACTGGTGTGTCAGCGCGAAGCTGAGCGTCAACAAAGACACACAGCTGACACTCGAGC 2445
Qy 927 TCGCGAGACCTCATGGCGGCGCGGCTGTGCGGCGCGCGCTTCAACGCGGCGCAAGGA 986
Db 2446 ACGCACCAACGAAGCGCGGACATCAGCTGCGGACCGCCAAAGCGGCTCTCGC 2505
Qy 987 GATCGCGCTGATCATTTTCATCGCGGACGATCGGTTTCTGACATATGCGGCGGCTG 1046
Db 2506 GACCATCGCTACACTACGACACGCGGCGGATCTGTCGAAGCGCGCGGCGCGCGGA 2565
Qy 1047 CGAGGAGAGCTCGCGGCGTTTCG 1071
Db 2566 CGCGCGGGTTCGCGGGGACCTCG 2590

RESULT 13
US-08-242-677-1
; Sequence 1, Application US/08242677
; Patent No. 5677143
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B
; APPLICANT: Wu, Foon W.
; TITLE OF INVENTION: Cellular Nucleic Acid Binding Protein
; TITLE OF INVENTION: and Uses Thereof in regulating Gene Expression and in the
; TITLE OF INVENTION: Treatment of AIDS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,677
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Mayfield, Denise L.
; REGISTRATION NUMBER: 33,732
; REFERENCE/DOCKET NUMBER: UTSD:401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5173 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4863
US-08-242-677-1

Query Match 6.1%; Score 71.6; DB 1; Length 5173;
Best Local Similarity 48.9%; Pred. No. 2.8e-05;
Matches 222; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Qy 158 CGAGCTTCCGTCACCGCGATGAGGACTGCTCCCGCGCTGCGGAGGCGCTGGCGG 217
Db 341 CGAGCTGCGCGCGCGCTGCTGAGAGGCGCTGCTGCGCGGATCTGCTCGCGGCTGGCGG 400
Qy 218 GGACGGGACACGAGTCCGCGGCGGCGCAGGCCATCGCCGAGCGCGCGGCGCGG 277
Db 401 GCCTGGCGCGGAGCTGCCGTGGAAGTGTAGCAGCGCTCGGGCCATGTTTTCGCGCGCC 460
Qy 278 CC---GAGTGAACCATGTCGGCCATCTGTGGGCGCGCGGCGCTGTCGCCGGAATCGCGG 334
Db 461 CGAGGACGCGCGCTACTGAGCGGGTGGCGGACCGCGCTGCGCTGCGCTGGCGCTGGCGG 520
Qy 335 CGCTGAACGACGCGCGCGCTGCGGCGCTGCGCCAAAGGAAGACCTCGTGACGCGGAGGC 394
Db 521 GGGCGGGGACGGGATGAGCGCGGCGCTGCGGAGAGCGCGCGCGCTGTGTGGCGCGGC 580
Qy 395 AACTCTGATGCGGACGCGCGCAGAGAACGCGCGCCACCATCTGCGCGTGGACAGCAGC 454
Db 581 GACTCTCCAGTGTGTCCATGTCGGGCGGCGCGCTGCGGGCGCGCTGTGGGCGCGG 640
Qy 455 ACTCGCGGCTTTTCAGGCGCTGCGGCGGCGGAGGACACGCGCTGCTGCGAGCGCGTCA 514
Db 641 TGGCGCGCGCTGGGCGCTCCCTGGGCTCGCGCGCTAGAGGAAGCTGCTGTGCTCTGA 700
Qy 515 TCAGGCGCTCGCGGCGCGCTTCCGCGACTGGAGCTCGAGCGCATCCGCGCGCTGCACG 574
Db 701 GCGCGCTGCGCGGAGGCTGTTGCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGG 760
Qy 575 TGGCGGAGCGCGAGCGCCCATCCCAACTGTGCTCAT 608
Db 761 AGCGGGCGCGGACGCGCGCGCGCTGCTGGCGCTT 794

RESULT 14
US-07-642-734C-3
; Sequence 3, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; TITLE OF INVENTION: Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
```

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dauckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20235 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238
FEATURE:
NAME/KEY: CDS
LOCATION: 19..10722
OTHER INFORMATION: /codon_start= 19
OTHER INFORMATION: /function= "gene eryA"
OTHER INFORMATION: /product= "eryA ORE2 encoding modules 3 & 4 for
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..4470
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 97..1482
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1693..2670
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3406..3921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4171..4428
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..10722
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..5847
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPsynthase domain of module"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19149..19398

LOCATION: 6054..7026
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7165..9216
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: dehydratase and enoylreductase domains m"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9433..9984
OTHER INFORMATION: /function= "approximate span
OTHER INFORMATION: beta-ketoreductase of module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10225..10483
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 4"
FEATURE:
NAME/KEY: CDS
LOCATION: 10723..20235
OTHER INFORMATION: /codon_start= 10723
OTHER INFORMATION: /function= "gene -eryA"
OTHER INFORMATION: /product= "Oxf3 encoding modules 5 & 6
OTHER INFORMATION: 6-deoxyerythronolide B formatio"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10723..15165
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10831..12174
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPsynthase domain of modul"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12379..13350
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14062..14610
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14857..15114
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15166..20235
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15172..16569
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPsynthase domain of modul"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16768..17721
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 18379..18921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19149..19398


```
; NAME/KEY: misc_feature
; LOCATION: 4471..5847
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of module"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6054..7026
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 4"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7165..9216
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: dehydratase and enoylreductase domains m"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9433..9984
; OTHER INFORMATION: /function= "approximate span
; OTHER INFORMATION: beta-ketoreductase of module 4"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10225..10483
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 4"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10723..20235
; OTHER INFORMATION: /codon_start= 10723
; OTHER INFORMATION: /function= "gene-eryA"
; OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
; OTHER INFORMATION: 6-deoxyerythronolide B formatio"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10723..15165
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10831..12174
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12379..13350
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14062..14610
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14857..15114
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15166..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15172..16569
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16768..17721
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: 18379..18921
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19149..19398
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19492..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: thioesterase domain of module 6"
; FEATURE:
; US-08-439-009A-3

Query Match      5.9%; Score 69.4; DB 3; Length 20235;
Best Local Similarity 43.7%; Pred. No. 7.4e-05;
Matches 304; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

QY 381 CGTGACGGCAGGCAACTCTGTATGCGGACGGCCCGCCAGGAGAACGGCGCCACGATCCTGCC 440
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 5829 CGTGGTTCGGGGCCAGGCCCGCGCTCGCCGAGACCTCGCGGACACACCCCGAGCTCGG 5888

QY 441 GGTGGACAGGACACTCCGCGCTCTTTAGGCGCTGGCGGGGAGAGACACGCGCTGCGT 500
      || || || || || || || || || || || || || || || || || || || || ||
Db 5889 CCTGACCGACGGCGGCTGGACGCTCGCGACCGCGAGGGCGGTTGACGTTGGAGCCGC 5948

QY 501 CGAGCGCGTCATCATCACGCGCTCCGCGGGCGCTTCCGCGACTGGAGCCCTCGAGCGCAT 560
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 5949 CGTGCTCGGGAGACACCGCGGGCGTGTGCGGGAGCTGGAGCGCGTGGCCGAGGGCGG 6008

QY 561 CCGCGCTCGACCGTGGCGGAGCGCAGGCCCATCCCAACTGCTTCCATGGGCGCAGCGGAT 620
      || || || || || || || || || || || || || || || || || || || || ||
Db 6009 CCGTCCGGCGCGCGCTCGCGCGGTGACCTCCGCGCGCGCAAGCCGCTCTCTGTGCTT 6068

QY 621 CTCATCGACAGCGCTCGATGTTTCAACAGGCGCTCGAGCTGATCGAGACGCGCGAATT 680
      || || || || || || || || || || || || || || || || || || || || ||
Db 6069 CCGCGCCAGGGCGCGACGTGGGTGCGCATGGCACGCGATCTGCTGGAATCTCTCCGAGT 6128

QY 681 CTTCGCTTCGAGCGGACCGGATCGAGGCGGTGCTTCCATCCGCAATCCATCTCCATGC 740
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 6129 GTTCGCGGAGTCGATGAGCGGTCGCGCGGAGCGCTCTCGCGCCACACCCGACTGGAATT 6188

QY 741 GATGGTGGGCTTCGCGAGGGGGCTGTATGCGCCCATCTCGGCGCCCGCCGACATCGGCCA 800
      || || || || || || || || || || || || || || || || || || || || ||
Db 6189 GCTCGACGTCGTCGCGGGGACGCGGTGCCGACCGCGACGAGCGGTGACGCTGCTCCA 6248

QY 801 CGCATCGGATTCGCGCTGNACTGGCGGGTCCGCGGAGTCCCGTCCCGCGATCGA 860
      || || || || || || || || || || || || || || || || || || || || ||
Db 6249 GCCGGTGCTTCTTCGATCATGTTCTCGTGGCGCGAGCTGTGGCGCGCGACCGCGGTGAC 6308

QY 861 CCTCGCACAGATTGCGAGCGCTCACTTCCAGAACGCTGACGAGCAACGCTTTCGCGCCCT 920
      || || || || || || || || || || || || || || || || || || || || ||
Db 6309 CCGCGCGCGCTGTCGCGCCACTCGAGGCGGAGATCGCGCGCGCACGTGGCGGGCGC 6368

QY 921 GAGGCTTGGCGGAGAGCTGATGGCGGGCGCGGCTGTGCGGGCGCGCCCTTCAACCGGGC 980
      || || || || || || || || || || || || || || || || || || || || ||
Db 6369 GCTGTCGCTGGAAGCGCGCGAAGTGTGGTGGCCCTCGCGCAGCCAGGTGTGCGCGAGCT 6428

QY 981 CRAAGGAGATCGCGCTCGATCATTTTCATCCGCGGACGATCGGCTTTCGGACATGGCGG 1040
      || || || || || || || || || || || || || || || || || || || || ||
Db 6429 CGACGACCGAGGGCGGATGTTGTCGGTGGCGCGTCCCGCGAGCTGGAGACCGGTGCT 6488

QY 1041 GGTGGTCGAGGAGACGCTCGCGGGCGCTTTCGACCG 1075
      || || || || || || || || || || || || || || || || || || || || ||
Db 6489 CGCGCGCTGGAGCGCGCTGTCGCGGTGGCGCGCG 6523
```

Search completed: April 22, 2003, 04:38:44
Job time : 4312 secs

STIC-Biotech/ChemLib

CRFF

9/4/25

4/15 01+02

From: Saidha, Tekchand
Sent: Sunday, April 13, 2003 10:11 AM
To: STIC-Biotech/ChemLib
Subject: 09/673198 - sequence search request

09/673198

Please search the data base and interference files for SEQ ID NOS : 30 and 31.

Thank you,

Jekchand Saidha

Primary Examiner

Art Unit 1652, CM1, Room No. 10D05

Mail Box 10D01

(703) 305-6595

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

30-394 aa
31-1182 na

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/15
Date Completed: 4/22
Searcher Prep/Review: 12
Clerical: _____
Online time: 12

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: es
WWW/Internet: _____
Other (specify): _____

